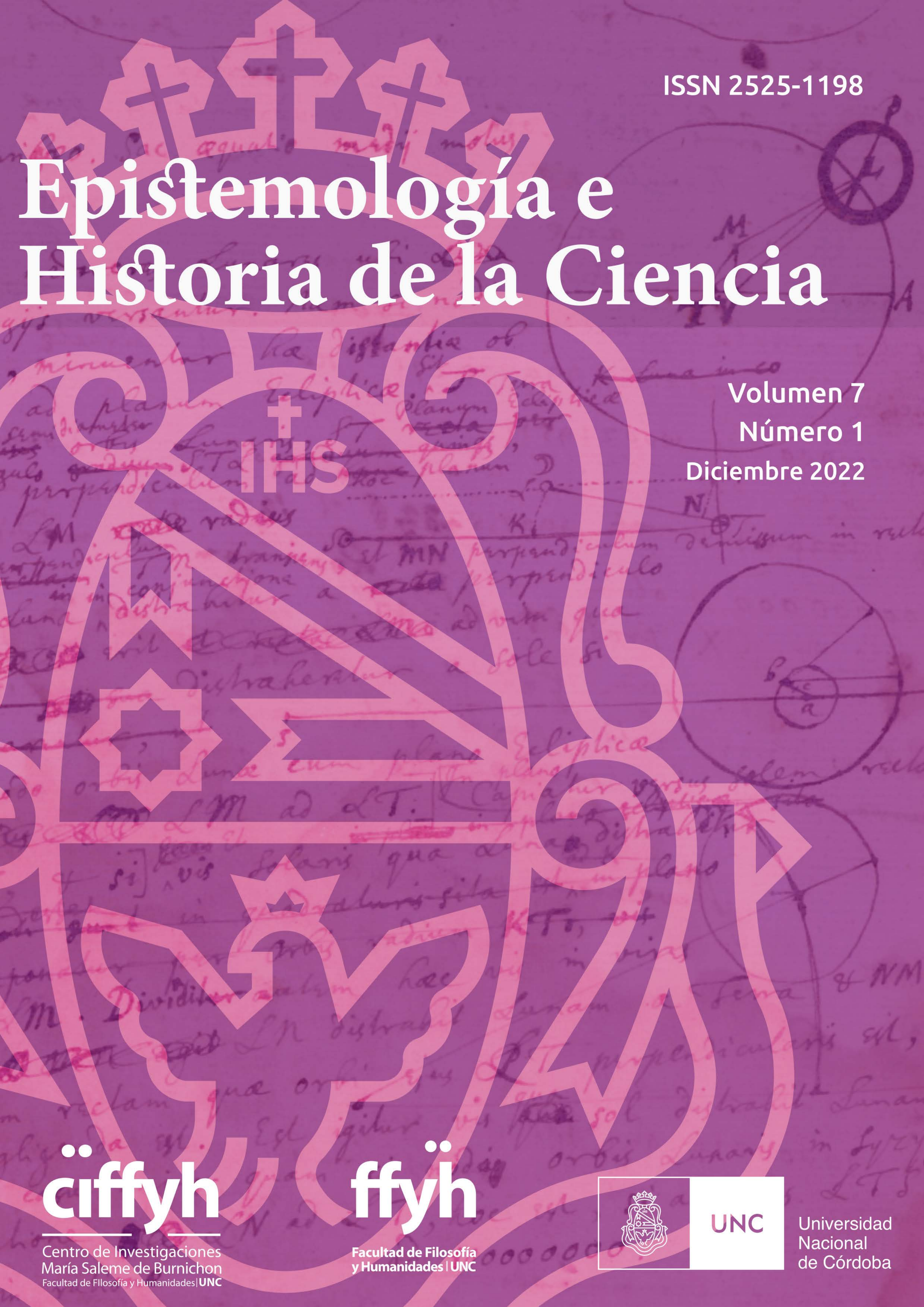


ISSN 2525-1198
Epistemología e Historia de la Ciencia
Volumen 7
Número 1
Diciembre 2022
ciffyh
Centro de Investigaciones María Saleme de Burnichon
Facultad de Filosofía y Humanidades | UNC
ffyh
Facultad de Filosofía y Humanidades | UNC
UNC
Universidad Nacional de Córdoba

Organismo Responsable

Área de Filosofía del Centro de Investigaciones de la Facultad de Filosofía y Humanidades y Escuela de Filosofía de la Universidad Nacional de Córdoba.

Pabellón Agustín Tosco, Ciudad Universitaria, Córdoba, Argentina.

Indexación

Catálogo Latindex 2.0 — Directory of Open Access Journals (DOAJ) — Sherpa Romeo — ERIH PLUS

ISSN: 2525-1198

Epistemología e Historia de la Ciencia

Epistemología e Historia de la Ciencia es una revista digital, de aparición semestral, dedicada a la publicación de artículos originales de filosofía general de la ciencia y filosofías de las ciencias particulares, así como artículos de historia de la ciencia con orientación filosófica. Las áreas de interés son entendidas en un sentido amplio y teóricamente plural.

Editorial y correspondencia:
Revista *Epistemología e Historia de la Ciencia*
Centro de Investigaciones Facultad de Filosofía y Humanidades (CIFFyH), Pabellón Agustín Tosco, Ciudad Universitaria, Córdoba (5000), Argentina.

Información adicional y envío de artículos:
https://revistas.unc.edu.ar/index.php/afjor
Correo electrónico:
revistaepistemologia@ffyh.unc.edu.ar

Epistemología e Historia de la Ciencia

Vol. 7 (Núm. 1)

Diciembre de 2022

Editorial

Artículos

Traducciones

In memoriam

Luis Salvatico (1963–2022)

Si para todo hay término y hay tasa
y última vez y nunca más y olvido
¿Quién nos dirá de quién, en esta casa,
sin saberlo, nos hemos despedido?
– Jorge Luis Borges

El pasado 24 de diciembre de 2022, de manera inesperada y repentina, falleció Luis Hobey Salvatico, quien fuera parte del Comité Editorial de *Epistemología e Historia de la Ciencia*, y con quien compartimos el lanzamiento de esta revista.

Luis era oriundo de la ciudad de San Francisco, Córdoba, Argentina. Allí cursó sus estudios primarios y secundarios, y luego realizó sus estudios universitarios de grado en la Universidad Nacional de Córdoba, Argentina, y su doctorado en la Universidad Nacional de La Plata, Buenos Aires, Argentina. Fue docente en distintas unidades académicas de la Universidad Nacional de Córdoba, donde también realizó su trayectoria profesional como investigador.

Epistemología e Historia de la Ciencia quiere recordar su aporte personal, su dedicación y su compañerismo, su constante compromiso con las Jornadas de Epistemología e Historia de la Ciencia, desde el inicio de las mismas, y de las cuales surge la iniciativa de lanzar esta revista. Podemos decir que Luis será recordado en cada espacio de nuestra vida académica; e incluso mucho más que eso, será recordado al sentir el dolor por un amigo que ya no está.

Un extractor de jugo teórico

El papel de las matemáticas en la explicación científica

Manuel Barrantes[1]

Recibido: 6 de septiembre de 2022
Aceptado: 7 de noviembre de 2022

Resumen: Recientemente se han propuesto casos en los que, supuestamente, las matemáticas cumplirían un rol genuinamente explicativo en ciencia. Estos se han dividido entre aquellas situaciones en las que el rol explicativo lo cumplirían las operaciones matemáticas, y aquellas en las que lo cumplirían las entidades matemáticas. En este artículo analizo algunos de estos supuestos casos y sostengo que es infundado afirmar que las matemáticas puedan ser genuinamente explicativas. En mi discusión, enfatizo en rol representacional de las matemáticas, opuesto al rol supuestamente explicativo: el rol de las matemáticas, incluso en los casos que discuto, es representar hechos científicos y ayudar a realizar inferencias acerca de dichos hechos.

Palabras clave: explicación científica, explicación matemática, representación, cigarras, arcoíris, panales.

Title: A theoretical juice extractor: The role of mathematics in scientific explanation

Abstract: There have recently been proposed cases where, supposedly, mathematics would play a genuinely explanatory role in science. These have been divided into those situations where the explanatory role would be played by mathematical operations, and those where it would be played by mathematical entities. In this article, I analyze some of these purported cases and argue that claims that mathematics can be genuinely explanatory are unfounded. Throughout my discussion, I emphasize the representational role of mathematics, as opposed to its supposed explanatory role: the role of mathematics, even in the cases that I discuss, is to represent physical facts and help draw inferences about those facts.

Keywords: scientific explanation, mathematical explanation, representation, cicadas, rainbows, honeycombs.

1. Introducción

La manera más natural de entender el rol de las matemáticas en explicaciones científicas es el de representar y ayudar a extraer inferencias sobre los hechos físicos responsables de la ocurrencia del fenómeno a explicar (el *explanandum*). De acuerdo a Joseph Melia, por ejemplo, dado que las matemáticas brindan buenas representaciones del mundo físico, algunas explicaciones científicas requieren el uso de matemáticas, pero esto no significa que las matemáticas sean en sí mismas explicativas. Si decimos, por ejemplo, que 'F ocurre porque P mide $\sqrt{2}$ metros', a pesar de que estamos mencionando el número $\sqrt{2}$ en la explicación, es el largo del objeto físico P, y no el número real $\sqrt{2}$ con el cual lo representamos, el que realiza el verdadero trabajo explicativo (cf. Melia, 2002, p. 76). La idea

[1] Departamento de Filosofía, California State University, Sacramento. Sacramento, California, Estados Unidos.
✉ barrantes@csus.edu | iD 0000-0002-7997-5948
Barrantes, M. (2022). Un extractor de jugo teórico: El papel de las matemáticas en la explicación científica. *Epistemología e Historia de la Ciencia, 7*(1), 6–21.
https://revistas.unc.edu.ar/index.php/afjor/article/view/38809/

es que las matemáticas aparecen en explicaciones científicas en virtud de dichos roles representacional e inferencial. Al derivar inferencias sobre la representación matemática, podemos aprender cómo los hechos físicos relevantes explican el *explanandum.*

Recientemente, sin embargo, se ha sostenido la tesis de que las matemáticas serían capaces de hacer más que esto, que podrían jugar un rol genuinamente explicativo en ciencia. Un supuesto tipo de dichos casos, propuesto por Robert Batterman, consiste en las explicaciones que incluyen las llamadas *operaciones* con límites. De acuerdo a Batterman, en estas explicaciones las matemáticas no representarían ningún aspecto de los sistemas físicos que intentan modelar, y sin embargo juegan un rol crucial en explicar aspectos de dichos sistemas. Por lo tanto, sostiene Batterman, el rol de las matemáticas en estos casos no es representacional sino explicativo en sí mismo. Otro tipo de casos, propuesto por Alan Baker (2005) y Aidan Lyon & Mark Colyvan (2008), son aquellos en los que *entidades* matemáticas serían esenciales para la explicación, de manera tal que, si no se mencionaran propiedades específicas de dichas entidades, la explicación no funcionaría. El propósito de este artículo es analizar estos casos y mostrar que es infundado afirmar que las matemáticas juegan un rol explicativo. El rol de las matemáticas, sostengo, es representacional incluso en los casos que usan operaciones con límites o que dependen de manera indispensable de propiedades matemáticas específicas.

El plan de este artículo es el siguiente. En la sección 2 introduzco la concepción inferencial de la aplicabilidad de las matemáticas, de acuerdo a la cual el rol de las matemáticas en explicaciones científicas es (simplemente) representar los hechos físicos que realizan el verdadero trabajo explicativo para poder extraer inferencias acerca de los mismos. También sostengo que esta postura es compatible con algunas de las más importantes posturas filosóficas sobre explicación científica, a saber, las de Carl Hempel, Wesley Salmon, y James Woodward. En la sección 3 introduzco la primera objeción a la postura inferencial, a saber, que algunas operaciones matemáticas son explicativamente indispensables a pesar de que no representan nada en el mundo físico. El ejemplo que discuto en esta sección es la supuesta explicación de dos aspectos relacionados de los arcoíris por la operación con límites que relaciona los modelos de la luz como un rayo y como una onda. En la sección 4 introduzco la segunda objeción, que consiste en que algunas explicaciones científicas apelarían de manera indispensable a entidades matemáticas, y por esa razón, dichas entidades estarían jugando un rol explicativo y no meramente representacional. Los casos que analizo aquí son el de los panales de abeja, y el de las cigarras. En la sección 5 sostengo, contra estas dos objeciones, que incluso en estos casos el rol de las matemáticas es representacional.

Antes de entrar en los argumentos, quisiera decir algo sobre la metodología que seguiré en adelante. Primero, he escogido los casos que voy a discutir porque son los más citados por los filósofos que cuestionan la postura representacional. Al mostrar que estos casos pueden ser explicados de otra manera, espero contribuir a defender la postura representacional. Sin embargo, mi argumento no pretende funcionar como algo que en principio aplicaría a todos los supuestos casos de explicaciones matemáticas en ciencia. Mi aproximación es 'por partes', digamos, en el sentido de que trabajo caso por caso. Habiendo dicho esto, sin embargo, sí considero que estos casos son representativos de las diferentes maneras en las que se puede lidiar con supuestos casos de explicaciones matemáticas en ciencia: se puede mostrar que una operación o entidad no es realmente indispensable para

la explicación, o que la explicación misma no es indispensable, debido a que existen mejores alternativas.

El segundo punto metodológico tiene que ver con haber elegido la concepción inferencial. En este artículo no defiendo esta postura como un todo. Simplemente sostengo que una de sus tesis más importantes, a saber, el carácter representacional de las matemáticas en explicación científica, es compatible con los modelos filosóficos más aceptados de explicación científica, y con los casos que supuestamente la cuestionan. Existen otras concepciones filosóficas de la aplicabilidad de las matemáticas en ciencia (por ejemplo, Suárez, 2010) que comparten este aspecto también.[2]

2. El rol de representacional de las matemáticas

De acuerdo a la Concepción Inferencial de la Aplicabilidad de las Matemáticas de Otávio Bueno, Steven French, y Mark Colyvan, en las explicaciones científicas que usan matemáticas, éstas deben interpretarse en términos del fenómeno objetivo. El formalismo matemático en sí no realiza el trabajo explicativo, sino los aspectos empíricos representados o resaltados por él (ver: Bueno & Colyvan, 2011; Bueno & French, 2012, 2018). En esta postura, las matemáticas pueden expresar relaciones explicativas solo cuando se les interpreta adecuadamente en términos empíricos:

> [Una] piedra es arrojada al aire. En cierto momento, la ecuación matemática que describe el movimiento de la piedra tiene valor cero. ¿Acaso el hecho de que la ecuación tenga dicho valor explica por qué la piedra está en reposo, o simplemente ofrece una descripción matemática del fenómeno en cuestión? Presumiblemente, nadie pensaría que el hecho de que una ecuación tenga valor cero sea en sí mismo una explicación de un fenómeno físico. Lo que se necesita para brindar una explicación satisfactoria es una adecuada interpretación física, que identifique los procesos físicos relevantes responsables de la producción del fenómeno en cuestión (Bueno & French, 2012, p. 104).

Esta manera de entender el rol de las matemáticas es compatible con algunos de los más importantes modelos filosóficos de explicación científica. Consideremos por ejemplo el modelo Nomológico-Deductivo (ND) de Carl Hempel. En el modelo ND, explicar científicamente consiste en deducir el *explanandum* de un conjunto de premisas que incluyan una ley empírica que sea esencial para la derivación. En esta postura, la herramienta explicativa relevante es la *deducción*, y las matemáticas son simplemente herramientas para llevar a cabo dichas deducciones. En palabras de Hempel, la función de las matemáticas es:

> Hace[r] explícitos ciertos presupuestos o aseveraciones que están incluidas en el contenido de las premisas del argumento. [...] Tanto las matemáticas como el razonamiento lógico son técnicas conceptuales que hacen explícito lo que está implícito en un conjunto de premisas. (Hempel, 1983, p. 390)

> Así, en el establecimiento de conocimiento empírico, las matemáticas [...] tienen, por así decirlo, la función de un extractor de jugo teórico. Las técnicas de teoría matemática

[2] Agradezco a los dos jueces anónimos de esta revista por la sugerencia de extender la introducción del artículo y añadir esta nota metodológica.

> o lógica no pueden producir más jugo de información factual que la que está contenida en los presupuestos sobre las que se aplican. (Hempel, 1983, p. 391)

En el caso del modelo Causal-Mecánico (CM) de Wesley Salmon, solo procesos e interacciones causales pueden ser explicativos; ya que las entidades matemáticas no tienen poderes causales en el sentido de Salmon (como procesos que transfieren una cantidad conservativa tal como la describen las leyes conservativas), el rol de las matemáticas no puede ser el de explicar el *explanandum*. Las fórmulas matemáticas citadas en el *explanans* se usan simplemente para representar dichos procesos e interacciones causales (cf. Salmon, 1984).

En el caso del modelo intervencionista de James Woodward (2003), explicar científicamente requiere información sobre las relaciones invariantes de dependencia responsables por la ocurrencia del *explanandum*. Una relación invariante de dependencia es aquella que se mantiene estable bajo intervenciones, de manera que es posible manipular un elemento de la relación interviniendo en el otro. A veces es imposible llevar a cabo dichas intervenciones, y sin embargo podemos recabar esta información al analizar un modelo matemático que represente dichas relaciones, por supuesto siempre y cuando tengamos un modelo apropiado. Así, por ejemplo, podemos explicar las mareas citando la posición de la luna en un momento determinado. A pesar de que intervenir sobre la posición de la luna no es posible, podemos manipular las mareas (es decir, podemos responder a preguntas del tipo 'qué-habría-pasado-si-las-cosas-hubieran-sido-diferentes') si 'cambiamos' la posición de la luna en el modelo matemático que la representa. Esto puede hacerse, por ejemplo, si insertamos un valor diferente en la variable que representa la distancia de la luna a la tierra, y vemos cómo la representación matemática de las mareas varía correspondientemente. Esto nos permite identificar la distancia de la luna a la tierra como una de las causas de las mareas. Aquí, el modelo matemático es simplemente una herramienta para manipular relaciones objetivas de dependencia (cf. Woodward, 2003, p. 129).[3]

La idea es, por lo tanto, que estos modelos filosóficos de explicación científica son compatibles con la postura de que el rol de las matemáticas es meramente representacional. Sin embargo, esta noción ha sido recientemente cuestionada por algunos autores que sostienen que las matemáticas cumplen un rol explicativo, en algunos casos adicionalmente a su rol representacional, y en otros en lugar de dicho rol. Estos casos han sido llamados 'Explicaciones Matemáticas de Fenómenos Físicos' (EMFF). Existen dos tipos de EMFFs en la literatura, las que apelan a *operaciones* matemáticas, y las que apelan a *entidades* matemáticas.

3. Explicaciones matemáticas: Operaciones

Robert Batterman (2010) sostiene que las matemáticas pueden cumplir un rol genuinamente explicativo, uno que es independiente de su rol representacional. Batterman introduce la noción de 'explicación asintótica', que es aquella que involucra de manera indispensable una operación matemática llamada razonamiento asintótico, que permite transitar entre dos modelos teóricos que convergen en el límite matemático. Lo interesante es que estos dos modelos representan el mismo fenómeno, pero sus presupuestos teóricos subyacentes son

---

[3] Una discusión más detallada de los modelos de explicación causales como el de Woodward y Salmon puede encontrarse en el artículo de Olga Gómez & Germán Guerrero en esta revista (Gomez Gutierrez & Guerrero Pino, 2020).

diferentes, y generalmente solo uno de ellos es tomado como correcto por la comunidad científica. Por ejemplo, el modelo A intenta ser una representación realista de un fenómeno físico, y el modelo B es una representación más idealizada del mismo. Sin embargo, el modelo B resulta de aplicar una operación asintótica al modelo A, en concreto, de llevar al límite a una variable del modelo A. Para Batterman, lo que es importante en estos casos es que la *operación* matemática (en este caso, obtener un límite) que relaciona ambos modelos es esencial para usar el modelo más realista en explicaciones de fenómenos físicos. Según Batterman, las explicaciones asintóticas no encajarían bajo los modelos representacionales de la aplicación de las matemáticas porque ellas "no proceden enfocándose en una estructura abstracta ejemplificada por el sistema físico" (2010, p. 3), y "*no* requieren que uno asocie una entidad matemática o sus propiedades con una estructura física que el sistema en cuestión posea" (2010, p. 4).

3.1. El caso de los arcoíris

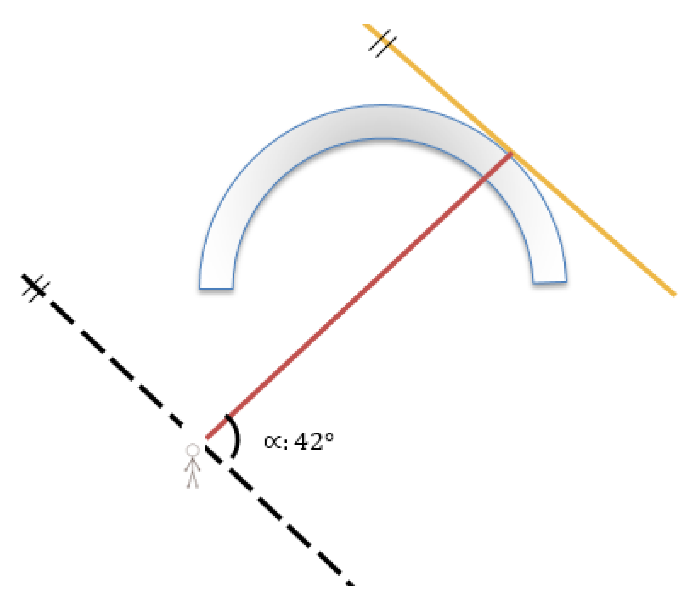


Figura 1: Ángulo de elevación de un arcoíris

Examinemos uno de los casos más citados en la literatura. Los arcoíris aparecen siempre con un ángulo de elevación de 42° relativo a la dirección de la luz solar (es decir que el ángulo que forman los rayos de luz que vienen de la parte superior del arcoíris con los rayos de luz del sol es aproximadamente 42° (ver figura 1). Adicionalmente, los arcoíris siempre aparecen con el mismo patrón de colores, a pesar de que cada arcoíris es el resultado de un conjunto único de circunstancias (Batterman, 2010, 20). ¿Cómo explicar ambas características?

Estas características de los arcoíris se deben a la interacción entre las ondas de luz y las gotas de lluvia, por lo que es natural pensar que se debe usar el modelo de luz como onda para explicarlos. Y, ciertamente, usando este modelo uno puede explicar el patrón de colores, como consecuencia del fenómeno de dispersión, a saber, que el índice de refracción de la luz varía cuando la luz pasa de un medio a otro (en este caso, los medios son el aire y el agua de las gotas de lluvia), dependiendo de la longitud de onda de las ondas de luz (los valores aproximados del índice de refracción en los extremos opuestos del espectro son $n_{rojo} \approx 4/3$ y $n_{violeta} \approx 1{,}344$.). Esto a su vez influye en el orden en el que las ondas inciden en el ojo

del observador, desde las ondas rojas provenientes de las gotas más altas, hasta las ondas violeta provenientes de las gotas más bajas, tal como puede observarse en la figura 2:

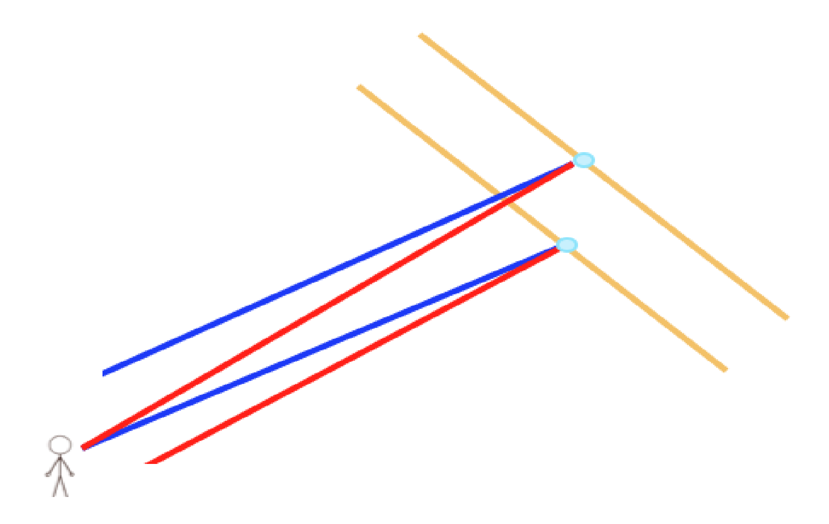


Figura 2: Incidencia en un observador de dos rayos de luz de distinta longitud de onda.

Sin embargo, para explicar la inclinación misma (por qué el ángulo entre la luz refractada y la luz solar es siempre aproximadamente 42°) tenemos que apelar a las relaciones geométricas entre la luz incidente y refractada al interior de las gotas de lluvia (asumiendo gotas perfectamente esféricas). En el caso de las ondas de luz rojas, las relaciones geométricas son las siguientes:

Siguiendo la ley de Snell:

$$\operatorname{sen}(2\beta - \phi) = n \operatorname{sen}\beta \qquad (1)$$

La concentración mayor de rayos ocurre cuando la variación de ϕ respecto a β es cero, lo cual puede representarse matemáticamente como:

$$\frac{d\phi}{d\beta} = 0 \qquad (2)$$

Para la luz roja, $n_{\text{rojo}} = \frac{4}{3}$, por lo tanto, tenemos:

$$\operatorname{sen}(2\beta - \phi) = \frac{4}{3}\operatorname{sen}\beta$$

$$\phi = 2\beta - \operatorname{sen}^{-1}\left(\frac{4}{3}\operatorname{sen}\beta\right)$$

$$\frac{d}{d\beta}[\phi] = \frac{d}{d\beta}\left[2\beta - \operatorname{sen}^{-1}\left(\frac{4}{3}\operatorname{sen}\beta\right)\right]$$

Usando (1) tenemos:

$$0 = 2 - \frac{\frac{4}{3}\cos\beta}{\sqrt{1 - \frac{16}{9}\operatorname{sen}\beta}}$$

Ya que:

$$\cos^2\beta = 1 - \operatorname{sen}^2\beta,$$

tenemos que:

$$\operatorname{sen}\beta = \sqrt{\frac{5}{12}}$$

Remplazando β en (1) obtenemos:

$$\phi \approx 21°.$$

Por lo tanto:

$$2\phi \approx 42°.$$

Un razonamiento análogo nos da $2\phi \approx 40{,}5°$ para la luz violeta.

Esto es sorprendente. La explicación depende de las relaciones geométricas que ocurren al interior de la gota de lluvia (Figura 3), las cuales se capturan cuando la luz se representa como perfectamente linear, es decir, cuando se usa el modelo antiguo de la luz como un rayo, evitando entrar en los detalles reales de la naturaleza de la luz, que es una onda (o una serie de paquetes de onas). Por el contrario, debemos amplificar la escala a la que representamos la luz, y asumir que la luz es un rayo, y no una onda (o paquetes de onda), lo cual es falso. El problema, señala Batterman, es que esta falsa representación es esencial para explicar el ángulo de elevación de 42°.

Entonces, para explicar el patrón de colores, debemos apelar a las propiedades de la luz tomada como onda (dependiendo de la longitud de onda, el índice de refracción varía); pero para explicar el ángulo mismo, tenemos que apelar a las relaciones geométricas que ocurren al interior de la gota de agua, lo cual a su vez requiere que describamos a la luz como un rayo. ¿Por qué una representación falsa es esencial para explicar un aspecto real de los arcoíris? Y más importante aún, ¿por qué necesitamos dos modelos diferentes, que asumen ontologías diferentes, para explicar aspectos relacionados de un mismo fenómeno?

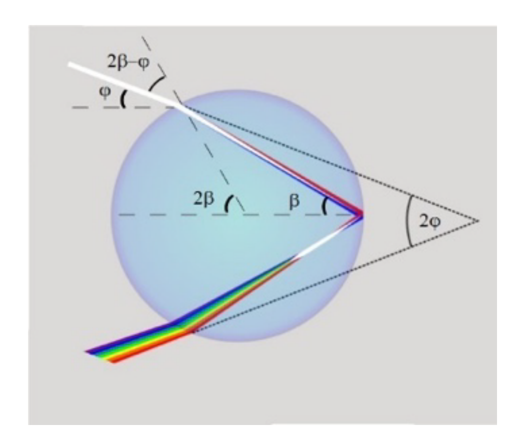


Figura 3: Comportamiento geométrico de la luz en una gota de lluvia perfectamente esférica. Imagen de dominio público, tomada de https://commons.wikimedia.org/wiki/File:Raindrop_optics.jpg

3.2. Operaciones límite

De acuerdo a Batterman, la razón por la que esta explicación tiene éxito es que el modelo de onda y el modelo de rayo están relacionados por una *operación asintótica*. En el modelo de onda, cada onda refractada tiene una longitud de onda asociada λ. Cuando la proporción entre esta longitud λ y el radio r de la gota de lluvia se aproxima a cero, la representación matemática realista colapsa en la representación falsa. Esta operación se llama 'tomar el límite'. Tomar el límite, dice Batterman, es explicativamente esencial porque provee "el

vínculo mediador entre los modelos representativos" (2010, 10). Una ilustración simple de esto es dada por una función f(x) que define una curva en un plano:

$$f(x) = A\cos\frac{2\pi}{\lambda}x \tag{3}$$

$$\lim_{1/\lambda\to\infty} f(x) = \lim_{1/\lambda\to\infty} A\cos\frac{2\pi}{\lambda}x$$

$$\lim_{1/\lambda\to\infty} f(x) = A \tag{4}$$

Como se puede ver, la ecuación (3) especifica una onda y la ecuación (4) especifica una línea. La operación de obtener el límite brinda una transición entre estas dos figuras, y debido a ello, sostiene Batterman, el modelo altamente idealizado del rayo luminoso puede ser usado exitosamente en la explicación de un fenómeno real, en este caso, el hecho de que el ángulo de elevación sea siempre el mismo. El modelo idealizado del rayo luminoso se obtiene como un caso límite del modelo más realista de la luz como onda (o paquete de ondas) como resultado de una operación matemática realizada sobre este último (la operación de obtener el límite), y en ese sentido, la operación en sí misma es explicativamente relevante.[4] La postura de explicación asintótica de Batterman se basa en el uso de las matemáticas para identificar patrones estables que omiten muchos detalles del fenómeno real. Al obtener el límite en el modelo matemático, uno ignora los detalles irrelevantes del proceso real.

De acuerdo a Batterman, estas explicaciones representan un problema para la concepción inferencial de Bueno, French, y Colyvan porque no se basan en similitudes estructurales entre el modelo matemático y el fenómeno físico objetivo. Batterman resume esta situación de la siguiente manera:

> El problema es simple. Nada en el mundo físico realmente corresponde con la idealización. Por lo tanto, ¿en qué sentido puede haber un mapeo desde una estructura matemática hasta una estructura física real? [L]as buenas representaciones reflejan la verdad acerca del mundo. Las idealizaciones, sin embargo, son falsas. (Batterman, 2010, p. 10)

En suma, Batterman sostiene que las explicaciones asintóticas representan un contraejemplo a las posturas representacionales, primero porque estas explicaciones dependen de los componentes idealizados de los modelos matemáticos, y segundo porque la operación misma no representa nada del mundo real. Batterman concluye que las posturas representacionales están equivocadas y que se necesita una postura completamente nueva para dar cuenta del rol que cumplen las matemáticas en explicaciones científicas.

4. Explicaciones matemáticas: entidades

En la literatura reciente ha habido mucha discusión sobre la posibilidad de explicaciones matemáticas en las que *entidades* matemáticas cumplan un rol crucial. La mayoría de las discusiones sobre EMFFs se han enfocado en estos casos.[5] En el contexto de los debates

---

[4] Esto es así a pesar de que el modelo de onda no es el más realista de los modelos de luz disponibles (el modelo más aceptado es el de los paquetes de onda, pero para los efectos de este artículo, todo lo dicho sobre el primero, se aplica al segundo también).

[5] Por una discusión detallada de explicaciones matemáticas en general, y EMFFs en particular (incluyendo el caso de las cigarras que discuto más abajo) ver Mary Leng (2005).

sobre el realismo matemático, estas explicaciones se han usado para justificar la existencia de objetos matemáticos, usando una nueva versión del Argumento de Indispensabilidad de W.O. Quine.[6] La idea es que, si tenemos razones para creer en la existencia de las entidades y procesos inobservables postulados por nuestras mejores explicaciones científicas (por ejemplo, electrones, campos electromagnéticos, procesos evolutivos, etc.), entonces también tenemos razones para creer en la existencia de las entidades matemáticas que figuran en las EMFFs. Voy a ignorar este debate en este artículo,[7] aunque mi argumento tiene consecuencias que al final son relevantes para dicho debate. Sostengo que incluso en las llamadas EMFFs, el rol de las matemáticas es (simplemente) representar las entidades y los procesos físicos responsables por el verdadero trabajo explicativo.

4.1. El caso de los panales de abeja

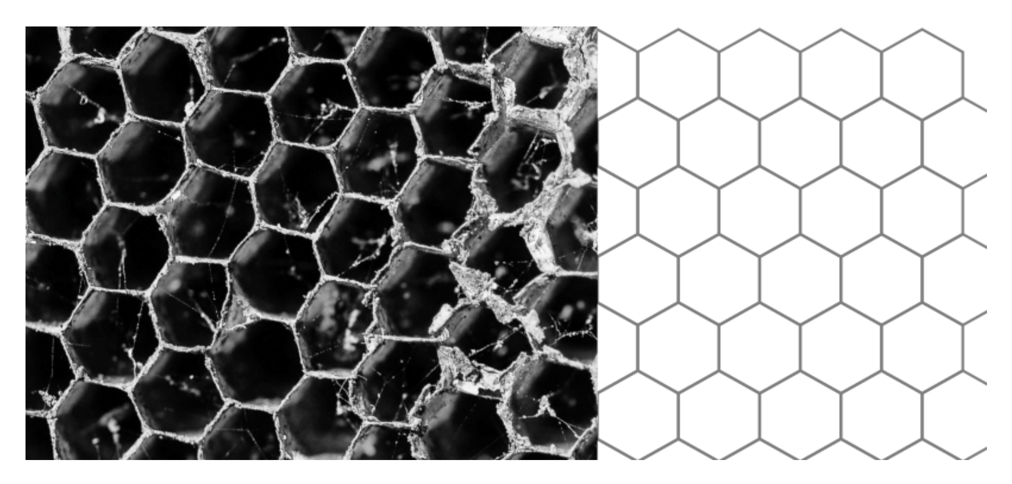


Figura 4. Hexágonos de un panel de abeja

Este caso fue propuesto por primera vez por Aidan Lyon & Mark Colyvan (2008) y ha sido ampliamente discutido en la literatura en el contexto del argumento de indispensabilidad. El problema es el siguiente:

Los panales de abeja dividen el espacio en mallas de celdas hexagonales en lugar de cualquier otro tipo de forma tales como triángulos, cuadrados, etc. (ver Figura 4). La explicación de este hecho constaría de dos partes:

> Es evolutivamente ventajoso para la las abejas minimizar la cantidad de cera que utilizan al construir sus panales para guardarmiel. Siguiendo el teorema de Hales (conocido como la ‘conjetura’ de los panales antes de ser probada por Thomas Hales en el 2001), la mejor forma de dividir una superficie en regiones de igual área minimizando el perímetro total es con una malla de celdas hexagonales. Por esa razón, las abejas construyen las celdas de sus panales de forma hexagonal

Esta explicación evolutiva depende del hecho matemático de que las mallas hexagonales tienen la propiedad de minimizar el perímetro total maximizando el área interior (teorema de Hales). La explicación muestra cómo, sin importar las diferentes formas que podrían haber intentado las abejas durante su historia evolutiva, una vez que

---

[6] El holismo confirmacional en el que se basaba la versión original ha sido criticado, entre otros, por Penelope Maddy (ver Maddy, 2005, por una síntesis de la discusión). La versión nueva se basa en el principio de inferencia a la mejor explicación (ver, por ejemplo, Baker, 2005). Para un análisis detallado de este mecanismo inferencial ver el artículo de Maximiliano Azcona en esta revista (Azcona, 2019).
[7] Abordo este debate en Barrantes (2019a) y (2019b).

intentan usar hexágonos este rasgo pasará a su descendencia, porque los hexágonos son evolutivamente ventajosos para cualquier abeja de construya panales. Según Lyon y Colyvan (2008), si no se menciona el teorema de Hales, que resalta una propiedad de una entidad matemática, la ventaja evolutiva de la forma hexagonal hubiera permanecido como un misterio. Por lo tanto, sostienen Lyon y Colyvan, este caso muestra cómo una propiedad de los hexágonos (que son entidades matemáticas) es indispensable para la explicación, y en ese sentido cumple un rol explicativo.

4.2. El caso de las cigarras

Este caso fue presentado por primera vez por Alan Baker (2005).[8] Existen siete especies de cigarras del género *Magicicada*, y estas emergen como adultas simultáneamente cada 13 o 17 años en Norteamérica, dependiendo de la región geográfica en la que se encuentren. Ciertas restricciones ecológicas brindan razones a los científicos para asumir que los ciclos vitales deben estar en rangos que incluyen dichos números. De acuerdo a Yin Yoshimura, por ejemplo, las cigarras del sur tendrían un rango de entre [12–15] años, y las del norte de entre [14–18] años (ver Yoshimura, 1997, pp. 113–114). Sin embargo, el hecho de que los ciclos vitales tengan un número de años primo no puede explicarse por estas restricciones ecológicas.[9]

Existen, sin embargo, dos explicaciones evolutivas disponibles. Una es que es evolutivamente ventajoso minimizar la posibilidad de intersección con depredadores; tener un ciclo vital primo minimiza las posibilidades de coincidir con depredadores de otros ciclos vitales. La otra es que es evolutivamente ventajoso evitar hibridación con otras subespecies de ciclos más largos o más cortos, pues eso llevaría a que los descendientes pierdan la ventaja de la emergencia simultánea; tener un ciclo vital primo reduce las posibilidades de dicha hibridación. Estas explicaciones evolutivas dependen de los siguientes hechos de teoría de números:

Un número es primo si admite solo un divisor aparte de 1. Dos o más números son co-primos si solo admiten a 1 como divisor común. En base a estas definiciones, se siguen los siguientes lemas:

Lema 1: el mínimo común múltiplo entre m y n es máximo sí, y solo sí, m y n son co-primos.

Lema 2: un número m es co-primo con cada número $n < 2m; n \neq m$ sí, y solo sí, m es primo.

Esto explica por qué, dado que algunas cigarras tienen ciclos vitales primos, tendrían a enfrentar menos depredadores que si tuvieran ciclos compuestos y estarían menos expuestas a otras subespecies con las que puedan coincidir. Por ejemplo, una subespecie de cigarra con un ciclo vital de 12 años se interceptaría con depredadores de ciclos vitales de 1, 2, 3, 4, 6 y 12 años, mientras que una subespecie con ciclo vital de 13 años solo se interceptaría con depredadores de ciclos vitales de 1 y 13 años. Un razonamiento similar se

[8] Discuto este caso en detalle en Barrantes (2019a).

[9] Contra Baker, Chris Daly y Simon Langford (2009) sostienen que la elección de años como unidad de medida es arbitraria, y por lo tanto los ciclos no son primos en sí mismos. Aunque estoy de acuerdo con esa conclusión (pues, como veremos, mi interpretación es que los intervalos temporales son atributos físicos, mientras que 'ser primo' es una propiedad de algunos números naturales), le doy la razón a Baker cuando responde que la unidad de medida no es arbitraria. La posición de la Tierra respecto del Sol afecta directamente los ciclos vitales de los organismos vivos al afectar aspectos locales relevantes tales como la temperatura y cantidad de luz diurna (ver Baker, 2009, p. 615). Las cigarras emergen en el verano, y solo existe un verano al año.

aplica a las cigarras de ciclos de 17 años. Adicionalmente, dado que ambas especies tienen ciclos primos, la indeseada hibridación entre cigarras con diferentes ciclos vitales ocurriría solo cuando los números de años de los ciclos se multiplican entre sí, es decir, cada 221 años.

El aspecto crucial de este ejemplo es que estas ventajas evolutivas se explican apelando a una propiedad matemática específica: la propiedad de ser primo, que comparten tanto 13 como 17. Tal como en el caso de los panales de abeja, si no se cita esta propiedad matemática, la explicación no funcionaría, y por esa razón, según Baker, estas entidades juegan un rol indispensable en la explicación. De acuerdo al biólogo Robert MacArthur, esta "sería la única aplicación de la teoría de números a la biología" (citado en May, 1979, p. 347).

5. Defensa de la postura representacional

En esta sección muestro cómo cada uno de estos casos puede ser comprendido bajo la postura representacional. Las matemáticas son una herramienta muy útil para realizar ciertas derivaciones, pero, finalmente, como señala Hempel, son un extractor de jugo teórico que ayuda a los científicos a extraer inferencias que identifican las verdaderas entidades y procesos explicativos, los cuales son, en todos los casos, empíricos. En estos tres casos que acabo de discutir, por lo tanto, las matemáticas no son indispensables en un sentido genuinamente explicativo.

En primer lugar, en el caso del arcoíris sigo a Bueno y French (2012), y a Pincock (2011), para sostener que la operación matemática en sí misma (la operación de calcular el límite) no contribuye a la explicación; solo los modelos matemáticos relacionados por ella lo hacen. Seguidamente, en el caso de los panales de abeja sostengo que existe una mejor explicación disponible, una que convierte a la original en obsoleta.[10] Si seguimos la idea de que una mala explicación no es una explicación en absoluto (de lo contrario, los astros explicarían la conducta humana), entonces las matemáticas no serían explicativas en el caso de los panales, tal como se pretendería en su formulación original. Finalmente, en el caso de las cigarras sí se cumple que la explicación dada sea la mejor disponible, pero el rol que juegan las matemáticas en ella es, después de todo, representacional, debido a que los ciclos vitales pueden ser interpretados como teniendo propiedades empíricas, no matemáticas.[11] Examinemos los tres casos en más detalle.

5.1. Respuesta al caso del arcoíris

De acuerdo a Bueno y French (2012), las explicaciones asintóticas pueden entenderse perfectamente bajo la concepción inferencial. Es cierto que, en principio, las operaciones límite no juegan un rol representacional, y esto podría señalar un problema para la concepción inferencial porque "si no hay análogos físicos que correspondan con las divergencias y singularidades del arreglo matemático, la concepción inferencial no puede interpretarse de vuelta en el arreglo físico" (Bueno & French, 2012, p. 91). Sin embargo, sostienen Bueno y French, siempre y cuando exista algún modelo que pueda interpretarse en términos del fenómeno físico objetivo, no hay problema en que este se relacione

[10] Estoy de acuerdo con un juez de esta revista en que este caso estrictamente no funciona como los otros, pues no estoy reinterpretando la explicación, sino que la remplazo por otra. He decidido conservarlo en el artículo porque me permite ilustrar una de las posibles maneras en la que se puede lidiar con supuestos casos de EMFFs.

[11] Esto también se aplica a la versión original del caso de los paneles de abejas, tal como lo muestra Saatsi (2011), aunque no voy a seguir esa línea argumentativa aquí.

estructuralmente con otro más idealizado. El punto es el siguiente: es cierto que no existe una interpretación física para las operaciones límite, pero si los resultados de dichas operaciones pueden interpretarse en términos del arreglo físico, la operación matemática en sí no tiene que por qué ser interpretada en términos físicos, pues esta se mantiene al nivel de una estructura excedente, es decir, como estructura matemática que cumple propósitos heurísticos (Bueno & French, 2012, p. 92).

Esto se puede ver claramente en el caso del arcoíris. Al obtener el límite ($\lambda/r \rightarrow 0$) estamos conectando el modelo de rayo con el modelo de onda, pero esto solamente muestra que el modelo de rayo es un caso especial del modelo de onda, y por lo tanto puede ser incorporado bajo el modelo de onda en situaciones en las que, una vez que el parámetro λ/r caiga por debajo de cierto umbral, algunos aspectos de las ondas son irrelevantes para explicar el ángulo de elevación del arcoíris. El punto de Bueno y French es que lo que hace el trabajo explicativo relevante en este caso es el modelo de rayo, y no la operación a través de la cual llegamos a dicho modelo.

Después de todo, incluso antes del advenimiento del modelo de onda de la luz, se podía considerar la explicación que describe a la luz como rayo como satisfactoria, una vez que se calculen empíricamente los índices de refracción. La manera como se llegue dicho modelo no es relevante para la explicación en sí. Pero la representación de luz como rayo sí tiene en cierto sentido una similitud estructural con la luz real porque, tal como señala Pincock (2011), esta resalta el hecho de que algunas relaciones al interior de la luz tienen un cierto tipo de 'linealidad', a saber, su trayectoria, la cual es esencial para explicar el ángulo de elevación. Si bien representar la luz como un rayo es incorrecto en varios sentidos, es correcto sin embargo respecto de este aspecto específico de la luz. Por lo tanto, a pesar de que no representa directamente el fenómeno objetivo, el modelo idealizado puede entenderse como una representación indirecta de algunos aspectos de dicho fenómeno.

5.2. Respuesta al caso de los panales de abejas

A pesar de la atención que ha recibido este caso, considero que el teorema de Hales no explica la forma hexagonal de los panales. Existe una explicación más tradicional que no apela al teorema de Hales, y en la que el rol de las matemáticas es evidentemente representacional. Fue propuesta por el ingeniero Bhushan Karihaloo, y reportada por Philip Ball en un número reciente de la revista *Nature* (Ball, 2013).

De acuerdo a Karihaloo, la forma hexagonal de los panales se debe tanto a las propiedades de la cera, como al procedimiento que siguen las abejas para construir sus panales. En un experimento, Karihaloo interrumpió a las abejas en el proceso de hacer los panales y encontró que, mientras que las celdas más antiguas eran hexagonales, las celdas nuevas eran circulares. Karihaloo concluyó que las abejas simplemente hacen celdas circulares agrupadas como una capa de burbujas. Debido a que las abejas calientan la cera con sus cuerpos mientras construyen las celdas circulares, una vez que terminan y se mueven a otra celda, la cera se enfría y se endurece en forma de hexágono (el hecho de que la cera caliente se retracta en hexágonos al enfriarse ya había sido probado por Christian Pirk en el 2004). Por lo tanto, reporta Ball, las celdas hexagonales "le deben más a simples fuerzas físicas que a la habilidad de las abejas" (Ball, 2013). Al representar estas fuerzas físicas tenemos que usar matemáticas, pero esto no significa que las matemáticas mismas jueguen un rol explicativo. Veamos en qué sentido esto es así.

La explicación esboza la historia causal del *explanandum*, mostrando que la forma hexagonal de las celdas se debe a procesos causales específicos de la cera. Se sugiere que la forma hexagonal no ocurre porque sea evolutivamente ventajosa, sino por las fuerzas producidas en la interacción entre la cera y el calor. Esta explicación se puede expresar en términos matemáticos, por supuesto, pero el rol que cumplen las matemáticas es muy parecido al que juega el número $\sqrt{2}$ cuando decimos que 'F ha ocurrido porque P mide $\sqrt{2}$ metros'. La explicación de Karihaloo socava la explicación que alude al teorema de Hales porque no es cierto que las abejas construyan celdas hexagonales (construyen celdas circulares que luego se vuelven hexagonales), por lo tanto, la explicación que se apoya en el teorema de Hales se basa en un presupuesto falso. Esto significa que la supuesta EMFF no explica la forma hexagonal.

Pero incluso la explicación original puede ser entendida en términos de las matemáticas representando fenómenos físicos, en este caso, relaciones geométricas en el espacio (esto ha sido argumentado en detalle en Saatsi, 2011). Esta estrategia puede verse mejor en el caso de las cigarras.

5.3. Respuesta al caso de las cigarras

En el caso de las cigarras, los ciclos de 13 y 17 años pueden entenderse perfectamente sin usar la noción 'primos'. Los ciclos poseen una propiedad física en común poseída por algunos intervalos temporales, que llamaré 'propiedad de minimizar la superposición'. Es cierto que hubiera sido muy difícil identificar esta propiedad sin la ayuda de las matemáticas, pero mientras sea posible concebir de manera separada la propiedad física de su representación matemática, el punto acerca del rol representacional de las matemáticas se sostiene. Y ciertamente podemos hacerlo en el caso de las cigarras. Las propiedades físicas del tiempo relevantes para este caso pueden deducirse a partir de las nociones básicas de *combinación* ($\oplus$) y *congruencia* de intervalos temporales, las cuales son propiedades no matemáticas perfectamente aceptables (por ejemplo, pueden representarse perfectamente en lógica de primer orden).

En el caso de las cigarras, la pregunta que intriga a los científicos es por qué los ciclos vitales representados por los números primos tienen una propiedad evolutivamente deseable. La representación matemática que asume que los años son iguales unos a otros, captura todos los factores relevantes para la explicación de este aspecto del tiempo. Al representar los años usando como modelo matemático al sistema de los números naturales, capturamos todos los aspectos relevantes para explicar por qué la propiedad de 'minimización de superposición' es evolutivamente ventajosa.

Es más, nos hemos basado en nuestros conocimientos previos acerca de la importancia de los ciclos naturales para explicar el comportamiento animal para escoger *años* como unidad de medida, y para idealizar los años como si fueran iguales unos a otros. Una vez que este modelo está establecido, nos podemos 'olvidar' de las cigarras y enfocarnos exclusivamente en el número primo de años. La pregunta ¿'por qué los ciclos vitales son evolutivamente ventajosos?' se convierte en 'por qué los números primos son co-primos con los demás?' La explicación matemática de este hecho responde la pregunta empírica sobre las cigarras solo en la medida en que nos muestra cómo las propiedades del tiempo explican los ciclos vitales. En otras palabras, usamos matemáticas para representar intervalos temporales, y luego usamos los lemas de teoría de números para entender por qué los ciclos son evolutivamente ventajosos. Cuando analizamos la representación

matemática, podemos ver que esto se debe a que los ciclos son primos; pero cuando interpretamos estos resultados en términos empíricos descubrimos que es la propiedad de minimizar la superposición la que explica la ventaja evolutiva. Es cierto que hubiera sido prácticamente imposible llegar a conocer dicha propiedad de los intervalos temporales sin usar la propiedad matemática de 'ser número primo', pero este es un tema simplemente de indispensabilidad práctica, y no de indispensabilidad explicativa. La propiedad matemática 'ser número primo' sirve simplemente para señalar la propiedad relevante del tiempo.

6. Conclusión

El tema de si las matemáticas pueden cumplir un rol genuinamente explicativo en ciencia es actualmente foco de mucha discusión tanto en la filosofía de la ciencia como en la filosofía de las matemáticas. En este artículo he mostrado que tanto en los casos que incluyen operaciones límite como en los que apelan de manera indispensable a entidades matemáticas, el rol de las matemáticas es representacional.

Agradecimientos

Agradezco a Valeria Panizo por sugerentes conversaciones que contribuyeron a mi enfoque en los arcoíris y las abejas, y a los jueces anónimos de esta revista por sus comentarios y recomendaciones. Una versión de este trabajo fue presentada en el Departamento de Matemáticas y Estadística de James Madison University, Virginia, USA.

Referencias

Azcona, M. (2019). Abducción e inferencia a la mejor explicación: criterios para su delimitación metodológica. *Epistemología e Historia de la Ciencia*, *4*(1), 33–55. https://revistas.unc.edu.ar/index.php/afjor/article/view/24955

Baker, A. (2005). Are there Genuine Mathematical Explanations of Physical Phenomena? *Mind*, 114, 223–238.

Baker, A. (2009). Mathematical Explanation in Science. *British Journal for the Philosophy of Science*, 60, 611–633.

Ball, P. (2013). How honeycombs can build themselves. *Nature*, Jul 2013. https://doi.org/10.1038/nature.2013.13398

Barrantes, M. (2019a), Optimal Representations and the Enhanced Indispensability Argument. *Synthese*, *196*(1), 247–263. https://link.springer.com/article/10.1007/s11229-017-1470-4

Barrantes, M. (2019b). Estructuralismo, ficcionalismo, y la aplicabilidad de las matemáticas en ciencia. *Areté. Revista de filosofía*, *31*(1), 7–34. https://doi.org/10.18800/arete.201901.001

Batterman, R. (2010). On the Explanatory Role of Mathematics in Empirical Science, *British Journal for the Philosiophy of Science*, *61*(1), 1–25.

Bueno, O., & Colyvan, M. (2011), An Inferential Conception of the Application of Mathematics, *Noûs*, *45*(2), 345–374.

Bueno, O., & French, S. (2012). Can Mathematics Explain Physical Phenomena? *British Journal for the Philosophy of Science*, *63*(1), 85–113.

Bueno, O., & French, S. (2018). *The Applicability of Mathematics: Immersion, Inference, Interpretation.* Oxford: Oxford University Press.

Daly, C., & Langford, S. (2009). Mathematical explanation and indispensability arguments. *Philosophical Quarterly 59*(237), 641–658.

Gomez Gutierrez, O. L., & Guerrero Pino, G. (2020). Explicaciones causal y mecanicista: aspectos etiológico, constitutivo, realista y pragmático. *Epistemología e Historia de la Ciencia*, *4*(2), 26–41. https://revistas.unc.edu.ar/index.php/afjor/article/view/25177

Hempel, C. (1983). On the nature of mathematical truth. En P. Benacerraf & H. Putnam (Eds.), *Philosophy of mathematics. Selected Readings* (2nd ed., pp. 377–393). Cambridge: Cambridge University Press.

Leng, M. (2005). Mathematical Explanation. En C. Cellucci & D. Gillies (Eds.), *Mathematical Reasoning, Heuristics and the Development of Mathematics* (pp. 167–189). London: King's College Publications.

Lyon, A., & Colyvan, M. (2008). The Explanatory Power of Phase Spaces. *Philosophia Mathematica*, *16*(2), 227–243.

Maddy, P. (2005). 'Three forms of Naturalism. En: S. Shapiro (Ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic* (pp. 437–459). Oxford: Oxford University Press.

May, R. M. (1979). Periodical cicadas. *Nature*, *277*, 347–349.

Melia, J. (2002). Response to Colyvan. *Mind*, *111*, 75–79.

Pincock, C. (2011), On Batterman's 'On the Explanatory Role of Mathematics in Science'. *British Journal for the Philosophy of Science*, *62*, 211–217.

Saatsi, J. (2011). The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science. *British Journal for the Philosophy of Science 62*(1), 143–154.

Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World.* Princeton: Princeton University Press.

Suárez, M. (2010). Scientific Representation. *Philosophy Compass*, *5*(1), 91–101.

Woodward, J. (2003). *Making Things Happen: A theory of Causal Explanation.* Oxford: Oxford University Press.

Yoshimura, J. (1997), The Evolutionary Origins of Periodical Cicadas During Ice Ages. *The American Naturalist*, *149*(1), 112–124.

Declaraciones

Conflictos de interés: El autor declara que no existen conflictos de interés.

La equivalencia formal en el lenguaje de las neuronas lógicas de McCulloch y Pitts

Rocío Stefanazzi Kondolf[1]

Recibido: 15 de agosto de 2021
Aceptado:17 de agosto de 2022

Resumen: En este artículo reflexionamos sobre el lugar de la equivalencia formal en el lenguaje propuesto por Warren McCulloch y Walter Pitts en su "A logical calculus of the ideas immanent in nervous activity" ["Un cálculo lógico de las ideas inmanentes en el sistema nervioso"]. Estudiamos el modelo a partir de los aportes que este ha significado en la historia de la ciencia: como teoría computacional de la mente; y como formalismo que contribuyó al desarrollo de la teoría de los autómatas y el diseño lógico. Estudiamos este *paper* desde la teoría de modelos y desde la teoría de la explicación mecanicista en ciencias. A su vez, reflexionamos acerca del carácter de cálculo lógico del texto en cuestión. Consideramos la relación entre el modelo y su *target* y ofrecemos dos interpretaciones posibles: las neuronas lógicas son el modelo que representa a la mente-cerebro o se trata de un nuevo objeto conceptual que posibilita el desarrollo de los autómatas artificiales. Observamos cómo la noción de equivalencia formal juega un papel fundamental: como explicación científica de una teoría de la mente, como criterio de isomorfismo de un modelo de la mente, y como criterio de identidad de las neuronas lógicas o neuronas formales.

Title: Formal equivalence in the language of McCulloch's and Pitts' logical neurons

Palabras clave: equivalencia formal, neuronas lógicas, nuevo necanicismo, modelos científicos.

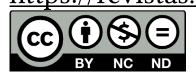
Abstract: In this article we consider the significance of formal equivalence in the language proposed by Warren McCulloch and Walter Pitts in their original "A logical calculus of the ideas immanent in nervous activity". We study the model from the contributions that it has meant in the history of science: as a computational theory of mind; and as a formalism that contributed to the development of automata theory and logical design. This question is framed into the mechanistic explanation account and the philosophy of scientific models, while also attending to how the model embedded a conception of a language as a logical calculus. We consider the relationship between the model and its target system, offering two possible interpretations: logical neurons are the model that represents the mind-brain or it is a new conceptual object that enables the development of artificial automata. We observe how the formal equivalence notion plays a fundamental role: as scientific explanation about mind's theory, as an isomorphism criterion of mind's model, and as identity criterion of logic neurons or formal neurons.

Keywords: formal equivalence, logical neurons, new mechanism, scientific models.

1. Introducción

En este texto reflexionamos sobre el rol de la equivalencia formal propuesta en el *paper* "A logical calculus of the ideas immanent in nervous activity" de Warren McCulloch y Walter Pitts escrito en 1943.[2] Resulta relevante reflexionar acerca de este *paper* por el lugar

[1] Universidad de Buenos Aires. Buenos Aires, Argentina.
[2] En este trabajo utilizamos la reedición de 1990 en el *Bulletin of Mathematical Biology* (McCulloch & Pitts, 1943/1990). El original se publicó en el *Bulletin of Mathematical Biophysics* (McCulloch & Pitts, 1943).

✉ rociosk3@gmail.com | 0000-0002-2755-0016
Stefannazi Kondolf, R. (2022). La equivalencia formal en el lenguaje de las neuronas lógicas de McCulloch y Pitts. *Epistemología e Historia de la Ciencia, 7*(1), 22–40.
https://revistas.unc.edu.ar/index.php/afjor/article/view/34428/


histórico que ocupa en la constitución de la cibernética (Abraham, 2002; Dupuy ,2000), su aporte a la teoría de la mente y sus influencias en el desarrollo de un lenguaje que contribuyó al surgimiento de las ciencias de la computación, los estudios de inteligencia artificial (IA) (Abraham, 2002; Piccinini, 2003,) y al desarrollo de las teorías del procesamiento de información (Abraham, 2002, pp. 3–4).

Este texto ha sido un puntapié para las investigaciones en la teoría de los autómatas y la arquitectura de von Neumann quien se inspiró en dicho trabajo en su "First Draft of a Report on EDVAC" (1945) y en "The General and Logical Theory of Automata" (1951). Algunos consideran que también sirvió de "modelo base" para la creación del *perceptron* de Frank Rosenblatt (Dupuy, 2000; Piccinini, 2003; Zamora-Cárdenas, Zumbado, Trejos-Zelaya, 2020, p. 17). En cuanto a la teoría de la mente, el modelo ha tenido influencia en el desarrollo de las ciencias cognitivas y las neurociencias (Abraham 2000).

El lenguaje que McCulloch y Pitts crearon ha sido leído por algunos autores, como es el caso de Gualtiero Piccinini (2004), como un mecanismo para explicar procesos de conocimiento y no el sistema nervioso. Dicho autor recorre las búsquedas teóricas y las experiencias de McCulloch y Pitts a los fines de comprender qué significó el trabajo de 1943 en tanto episodio de la historia de la ciencia. En este sentido, nos indica cuatro aportes: (1) un formalismo que permite la generalización de la noción de autómata finito, (2) una técnica inspirada en la noción de diseño lógico, (3) el primer uso de una computación de manera conjunta al problema mente-cuerpo, y (4) la primera teoría computacional de la mente y el cerebro. Que Piccinini (2004) considere que la propuesta del texto en cuestión sea un *mecanismo* para explicar el funcionamiento procesos mentales como el conocimiento u otros, está sostenido en la trayectoria de investigación de McCulloch, quien había trabajado en la teoría de los psicones. Los psicones eran algo así como "átomos mentales" equivalentes a proposiciones sobre sus antecedentes temporales. La teoría de los psicones fue clave para la resolución propuesta del problema mente-cuerpo: a cada átomo mental correspondía una actividad neuronal. A la equivalencia psicón-actividad neuronal se agrega la equivalencia con la actividad proposicional. El mecanismo de la mente es el mecanismo del cerebro y estos a su vez coinciden con el mecanismo de una lógica proposicional bivalente. A su vez, si el mecanismo puede ser considerado un mecanismo computacional, entonces para Piccinini (2003, 2004) la teoría de McCulloch y Pitts es una teoría computacional de la mente.

Por otro lado, autores como Mark Schlatter y Ken Aizawa (2008) ponen el foco en los aspectos formales-matemáticos, imbricados a la historia y las búsquedas teóricas de los investigadores, pero con más énfasis en la trayectoria de Pitts. En este sentido, consideran que McCulloch aportó la idea del uso del álgebra de Boole inspirado en el mecanismo "todo-nada" del modelo de neuronas biológicas. A esto Pitts lo conjugó con un lenguaje de funciones, que era un producto de sus trabajos en las redes neuronales de Alston Householder; y una escritura lógica de segundo orden, dada por el lenguaje II de Rudolf Carnap, con profunda influencia de Bertrand Russell. A esto que indican Schlatter y Aizawa, podríamos sumarle, además, una organización del texto mismo en términos de un cálculo o lenguaje lógico, probablemente también por la influencia de Carnap y, previamente, de Russell.

En este artículo retomamos los cuatro aportes del *paper* de McCulloch y Pitts que resume Piccinini (2004) mencionados anteriormente y, a partir de allí, evaluamos los sentidos de equivalencia formal, y el rol que juega esta noción en la articulación de cada

uno de dichos aportes. Utilizamos el marco de la explicación mecanicista para comprender el sentido de "mecanismo" que Piccinini pone de manifiesto en las teorizaciones de la mente de McCulloch. A su vez, utilizamos la teoría de modelos para pensar la relación representativa del modelo McCulloch-Pitts en cuanto al mecanismo de la mente y el diseño lógico de computadoras. Consideramos que si bien el modelo fue pensado para representar la mente-cerebro, en sus intentos de abstraer lo que concebían como su mecanismo lógico inmanente, posibilitaron la construcción de nuevas entidades conceptuales: las neuronas lógicas o formales. En este punto surge una cuestión doble, o las neuronas lógicas son el modelo de la mente-cerebro que sería el *target* u objetivo de aquel, o se trata entonces de un nuevo fenómeno *target* que posibilitó el desarrollo de la teoría de los autómatas artificiales.

Nos centramos en la primera parte del artículo de 1943 relativo a las redes neuronales sin círculos y mencionamos solo algunas cuestiones de la segunda parte, relativas a las redes con círculos. Reflexionamos sobre el lugar de la equivalencia formal en la constitución de las neuronas lógicas o neuronas formales del lenguaje propuesto por Warren McCulloch y Walter Pitts en su artículo. Observamos cómo la noción de equivalencia formal juega un papel fundamental: como explicación científica de una teoría de la mente, como criterio de isomorfismo de un modelo de la mente, como criterio de identidad de las neuronas lógicas. Insertamos esta cuestión en el marco de las explicaciones mecanicistas en filosofía de las ciencias, la reflexión contemporánea sobre modelos científicos y la concepción de los lenguajes como cálculos lógicos.

2. Los recorridos e influencias lógico-matemáticas

En este apartado desarrollamos los recorridos e influencias en el ámbito matemático, lógico y computacional Warren McCulloch y Walter Pitts mientras que en apartados posteriores nos dedicaremos a reconstruir las influencias de la neurofisiología.

Como nos indica Piccinini (2004), la formación y dedicación principal de McCulloch era la de psiquiatra y neurofisiólogo. Sus primeras investigaciones postulaban unas entidades que ya mencionamos: los "psicones"[3], una especie de átomos mentales que, luego asociados a expresiones proposicionales, le permitieron pensar una teoría de la mente y del conocimiento. Es en este contexto, según nos indica Piccinini, que McCulloch piensa la aplicación del álgebra booleana a las redes nerviosas.

En su búsqueda de una correlación lógica para sus átomos mentales, McCulloch se acercó al seminario filosófico en el cual Frederic Fitch dictó unas clases sobre los *Principia Mathematica* de Russell y Whitehead, en la Universidad de Yale. En este mismo marco, se acercó al trabajo de Nicolas Rashevsky (con quien trabajaba Pitts) a través Filmer Northrop —director del seminario filosófico— quien reivindicaba la importación de técnicas y formalismos de la física-matemática a la biología; en 1937, también en Yale, se acercó al trabajo de Joseph Henry Woodger acerca del uso de los sistemas axiomáticos en biología.[4] A

---

[3] El término en inglés es "psychons".
[4] Abraham (2002) rastrea en Northrop (1931): "La etapa descriptiva y clasificatoria en biología que comenzó con Aristóteles había comenzado a moverse hacia una etapa formal y deductiva con el trabajo de Joseph Henry Woodger (1894–1981) (Woodger, 1937) y Nicolas Rashevsky (1899–1972). En estos trabajos, la lógica formal y las matemáticas jugaron un papel importante, por lo que la biología como disciplina fue alcanzando una etapa más madura, en tanto que comenzó a incorporar el método científico de la física, es decir, utilizando el análisis teórico y las formulaciones matemáticas. Los argumentos de Northrop sobre el valor de la formalización en biología estaban conectados con su visión más amplia de 'diseccionar las teorías científicas dadas que. . . los científicos han verificado, para determinar qué conceptos y principios se toman como primarios o indefinidos' (Northrop, 1931, p. xiii)" (Abraham, 2002, pp. 6–7).

principios de 1940 conoce los trabajos de Alan Turing sobre las máquinas universales[5]. Un año después conocerá el trabajo de Rashevsky en biología matemática en la Universidad de Chicago y, este contexto, conocerá el trabajo de Pitts.

Schlatter y Aizawa (2008) nos indican que Pitts comenzó sus investigaciones en el grupo de Rashevsky en la Universidad de Chicago y, en particular, a partir de los trabajos de Householder (1941, 1942). La diferencia importante del *paper* de 1943 con el enfoque de Householder (1941a, 1941b, 1941c, 1942), se debió a que, en vez de trabajar con matemáticas continuas, lo hizo con matemáticas modulares o discretas.[6] Sin embargo, Pitts ya había introducido varios cambios a las explicaciones de Householder en una serie de artículos que había escrito previamente. Mientras este último había planteado la relación entre las neuronas en términos de sistemas de ecuaciones, Pitts planteó la estimulación en función del tiempo. De esta manera, resolvió el problema planteado por su antecesor: dada una red con un patrón de estimulación específico, encontrar el patrón de excitación. Otro cambio que produjo fue pensar en una sinapsis a lo largo del tiempo en vez de en la red en su totalidad. Exploró cómo cambia la estimulación de las sinapsis y, con el paso del tiempo, se acerca al estado estacionario [*steady-state*[7]]. A su vez, consideró otras posibles redes topológicas y distinguió entre sinapsis de 1er (una fibra de entrada y una de salida) y de 3er orden (una fibra de entrada y más de una de salida). Por último, para dar cuenta de los valores posteriores de estimulación, introdujo el operador ***E*** que se comporta de la siguiente manera:

- $\mathbf{E\,f(x) = f(x + 1)}$**:** El operador ***E*** remite a la estimulación una unidad de tiempo después de x
- $\mathbf{E^3\,f(x) = f\,(x + 3)}$**:** El operador ***E*** elevado a la tres remite a la estimulación tres unidades de tiempo después de x

Este operador es importante porque luego se utilizará uno semejante en el *paper* de 1943, llamado operador **S**.

Una segunda influencia muy importante para Pitts fue la de Rudolf Carnap a quien conoció por su interés en el trabajo de Russell y Whitehead. En su adolescencia había leído los *Principia Mathematica*, luego se acercó a unas conferencias que Russell dictó en Chicago, y éste lo contactó con Carnap. De Russell y de Carnap tomarán la idea de cálculo lógico, y en particular de Carnap una escritura lógica de segunda orden, dada por su lenguaje II.

---

[5] Hoy conocidas como máquinas de Turing. Tal como la define Turing (1936) una máquina de computar [*computing machine*] es aquella que tiene una serie finita de condiciones $q_1, q_2, \ldots, q_n$ las cuales se llaman m-configuraciones [*m-configurations*] y una cinta dividida en secciones [*squares*] las cuales pueden llevar un símbolo (0 ó 1). Una de esas secciones es el escáner de la máquina [*scanned square*] y el símbolo en dicha sección se llama símbolo escaneado [*scanned symbol*]. El símbolo escaneado es el único del cual la máquina es "directamente consciente", aunque si se cambian algunas configuraciones, ésta puede recordar algunos símbolos que ya han sido escaneados previamente. El comportamiento de la máquina queda determinado en cualquier momento por las configuraciones y el símbolo escaneado. Si la sección está en blanco la máquina puede escribir un símbolo y si tiene un símbolo puede borrarlo. La máquina solo cambia desde la posición que ha sido escaneada y siempre cambia de a un lugar moviéndose hacia la derecha o a la izquierda. Una máquina universal es aquella que puede computar cualquier secuencia computable que computen otras máquinas posibles (Turing, 1936, pp. 231, 232, 241, 242).

[6] Es decir, aquella que trabaja con funciones con cortes, que no son continuas, como es el caso de la lógica y el álgebra.

[7] "La actividad en estado estacionario [*steady-state*] bajo estimulación constante es la más simple de considerar [...]. Es decir, si se aplica un estímulo constante durante un período de tiempo a cada una de las fibras del complejo, lo que sea que las fibras hagan en los primeros milisegundos, o incluso segundos, suponemos que establece un breve estado de respuesta que persiste sin cambio mientras los estímulos se mantienen sin cambios" (Householder, 1941a, p. 64).

En síntesis, las influencias han sido varias: Por un lado, Northtrop, Woodger, Rashevsky y Householder en cuanto al uso de la matemática en biología; Boole, Russell y Carnap en cuanto a las influencias lógico-matemáticas y Turing en cuanto a la concepción de computabilidad.

3. Una teoría de la mente

En este apartado seguimos la reconstrucción que hace Piccinini (2003, 2004) de la teoría de la mente de McCulloch y Pitts teniendo en cuenta las siguientes cuestiones que fueron mencionadas anteriormente: que se trata de la primera teoría computacional de la mente; que es una teoría mecanicista de la mente; que necesitó de una concepción de entidades atómicas llamadas psicones —desarrollada por los trabajos anteriores de McCulloch— para establecer una equivalencia con la actividad neuronal; y que, de este modo, dio una respuesta al problema mente-cuerpo con la identificación de la mente y el cerebro.

3.1. La primera teoría computacional de la mente

Para Piccinini (2004) la primera teoría de la mente de McCulloch y Pitts es la primera teoría computacional en tanto utiliza la noción matemática de computación de Turing, incluso antes que éste último esbozara opiniones sobre la relación de su teoría computacional con la teoría de la mente (p. 176). Sin embargo, que hayan usado la noción de computación de Turing, no significa que las redes neuronales podrían computar cualquier cosa que compute una máquina de Turing, lo cual ha sido un error usual de interpretación (Piccinini, 2004, p. 177). Este error puede rastrearse en la influencia que McCulloch y Pitts tuvieron sobre la interpretación de la tesis Church-Turing. En la apelación a Turing que hacen en su texto aparece la siguiente idea:

> Se muestra fácilmente: primero, que cada red, si está equipada con una cinta, escáneres conectados a aferentes y eferentes adecuados para realizar las operaciones motoras necesarias, puede calcular sólo tales números como una máquina de Turing; segundo, que cada uno de los últimos números puede ser computado por tal red; y que las redes con círculos pueden ser computadas por tal red; y que las redes con círculos pueden computar, sin escáneres y una cinta, algunos de los números que la máquina puede, pero no otros, y no todos. Esto es de interés como una justificación psicológica de la definición de Turing de computabilidad y sus equivalentes, la definibilidad λ de Church y la recursividad primitiva de Kleene:[8] si cualquier número puede ser calculado por un organismo, es computable por definición, y de manera conversa. (McCulloch & Pitts, 1943/1990, p. 113)

Es decir, no están diciendo que sus redes sean matemáticamente equivalentes a la noción de computabilidad de Turing, como usualmente se ha comprendido erróneamente, sino que estas redes son una justificación psicológica de la posibilidad lógica de la tesis de Turing-Church. Tal como señala Piccinini (2003):

> La conexión trazada por McCulloch y Pitts, a través de sus redes, entre los fenómenos mentales y la computación ha tenido efecto en el modo que se interpretó la Tesis de Church-Turing (TC). Según la teoría de McCulloch y Pitts, cada red podría ser descripta como computando una función. Entonces, en el sentido en que las redes de McCulloch-

[8] Posteriormente Kleene (1956) retomó las cuestiones planteadas por McCulloch y Pitts y mostró que este formalismo era equivalente a los autómatas finitos.

> Pitts computaban, y en la medida en que las redes de McCulloch-Pitts eran un buen modelo del cerebro, cada actividad neuronal era una computación. Esto fue particularmente significativo dado que McCulloch y Pitts consideraban que las computaciones de sus redes eran explicaciones de los procesos mentales. (Piccinini, 2003, p. 63)

La teoría de McCulloch-Pitts parecía ser entonces una fundamentación psicológica de los límites de lo computable pero no así una tesis sobre los límites lógicos de lo computable.

3.2 Una teoría mecanicista de la mente

El llamado nuevo mecanicismo es una corriente heterogénea de pensadores y posturas en torno a la definición de mecanismo, pero comparten la reflexión común acerca de dicha noción. Las explicaciones mecanicistas surgen en el contexto de las llamadas *ciencias frágiles* (Barberis, en prensa) y no pretenden una explicación general de una teoría sino explicaciones locales y situadas de ciertos fenómenos en términos de mecanismos. Este tipo de explicaciones, como nos indican Branca, Ramírez, y Vilatta (2015), discuten con el reduccionismo que ha presentado a las explicaciones científicas como forma de reducir una teoría a otra. Craver (2007) ha indicado que la tradición reduccionista construye las explicaciones a través de enunciados de identidad en los que un nivel se deriva de otro y se requiere un mapeo uno-a-uno (pp. 107–108). En este sentido, la concepción mecanicista de McCulloch-Pitts establece enunciados de identidad que no son reduccionistas sino que todos los niveles involucrados en el mapeo uno-a-uno tienen la misma jerarquía.

En este punto seguimos el rastreo del mecanicismo que Piccinini (2004) hace de la teoría de la mente de McCulloch y Pitts a partir de las concepciones de los trabajos previos de McCulloch, para quien uno de los objetivos y aportes de la neurofisiología ha sido explicar la mente "en términos de mecanismos neuronales" (p. 177). En esta dirección, Dupuy (2000) nos dice que "McCulloch [...] fue un hombre poseído por una sola idea, la cual dependía de sostener al mismo tiempo cerebro, mente y máquina" (p. 111).

Tal como señalamos anteriormente, Piccinini (2004) encuentra en el propio discurso del neurofisiólogo una teoría de la mente que suponía átomos mentales —los llamados psicones—, las entidades mentales equivalentes a la actividad proposicional de las neuronas:

> [...] un psicón es ‘’equivalente’’ a una proposición sobre su antecedente temporal. En terminología más reciente, McCulloch parecía pensar que un psicón tiene un contenido proposicional, que contiene información sobre la causa de ese psicón. Un segundo punto clave fue que un psicón "propone" algo a un psicón posterior. Esto parece significar que el contenido de los psicones se puede transmitir de un psicón a otro, generando "los equivalentes" de proposiciones más complejas. (Piccinini, 2004, p. 178)

McCulloch comenzó a elaborar su teoría de los psicones a mediados de 1920. En 1928 hizo una estancia en Nueva York en el hospital Bellevue en el cual trabajó en la teoría de la función nerviosa y a partir de la cual pensó que ciertos problemas como la epilepsia podrían deberse a bucles o *loops* en la actividad nerviosa, y creía encontrar sustentos en las investigaciones de Ramón y Cajal. Una influencia importante en sus redes con *loops* fue el neurólogo y psicoanalista Lawrence Kubie (Dupuy, 2000, p. 55). Estos bucles o *loops*, asociados a problemas como la epilepsia o la enfermedad de Parkinson, reaparecerán en el artículo de 1943. En 1929 comienza a pensar en la hipótesis de que los impulsos eléctricos todo-nada de las neuronas podrían corresponderse a los psicones y que "las relaciones de excitación e inhibición entre las neuronas realizarían operaciones lógicas sobre señales

eléctricas correspondientes a inferencias de su cálculo proposicional de psicones.” (Piccinini, 2004, p. 179). Luego, en 1934, trabajó con Joannes Dusser de Barenne en su laboratorio de Neurofisiología en el mapeo de las conexiones entre áreas del cerebro. (Piccinini, 2004, pp. 178, 179, 180). Es en el discurso de la neurofisiología, en la equivalencia entre neuronas biológicas y psicones, que McCulloch encuentra la justificación de su concepción mecanicista de la mente y a partir de las preguntas que surgen en el seno de la neurofisiología incorpora nuevas disciplinas, en particular, la lógica, la matemática, la psicología.

En resumen, la estrategia de McCulloch y Pitts, según la lectura de Piccinini (2004) que venimos siguiendo, fue: (a) en primer lugar, reducir y simplificar las redes neuronales, y, de esta manera, una serie de inferencias proposicionales podían ser mapeadas en eventos neuronales y viceversa. En segundo lugar (b), asumir que los pulsos neuronales tenían contenido proposicional que se correspondían con procesos mentales atómicos. Esta teoría de la mente asociada a las expresiones proposicionales a través de la dinámica misma del impulso nervioso necesitó del supuesto mente = cerebro disolviendo de algún modo, o dando una respuesta al problema mente-cuerpo.

4. Un modelo de la mente-cerebro

Las teorizaciones de McCulloch y Pitts han producido un modelo que, a la hora de construir su *target*, tomó como punto de partida una serie de asunciones o supuestos “físicos”, tal como los nombran los propios autores del modelo. Potochnik, Colombo y Wright (2019) definen que “Una asunción […] es una especificación que un sistema *target* debe satisfacer para un modelo dado, de ser similar en el modo esperado”, y que “Las asunciones pueden ser idealizaciones que no necesariamente serán verdaderas.” (p. 99). Es decir, que un *target* no necesariamente se corresponderá con el fenómeno real. Aunque los autores mencionados anteriormente tienden a considerar al *target* como el fenómeno real (Potochnik, Colombo, & Wright, 2019, p. 96). En contraposición a este tipo de posturas realistas, Cassini (2018) considera que el *target* de un modelo no es el fenómeno real, sino el producto de un proceso de construcción científica. Lo que nos interesa principalmente en este apartado es rastrear las asunciones que se han abstraído e idealizado[9] en el modelo que nos posibilitarían rastrear el *target* del mismo.

4.1. La neurofisiología teórica y los “supuestos físicos”

En el artículo de 1943 McCulloch y Pitts toman el modelo de la neurona de lo que llaman la “neurofisiología teórica”, sobre la cual no citan ninguna bibliografía.[10] En la “Introducción” describen al sistema nervioso como una red de neuronas a las cuales le atribuyen dos grandes partes: el soma y el axón. Es notable que no toman en cuenta a las dendritas. El axón de una neurona se conecta con el soma de otra mediante la sinapsis, a través de impulsos nerviosos que pueden ser inhibitorios o excitatorios. Hay un período de adición latente en el que la neurona puede recibir impulsos. Para que se inicie un impulso nervioso, la excitación de la neurona debe exceder el límite o umbral que está determinado por la neurona misma. Una vez iniciado el impulso, se propaga desde el punto de excitación hacia

[9] “Omitir o ignorar ciertas características conocidas del sistema es una abstracción; incluir características que el sistema target no tiene es una idealización” (Potochnik, Colombo, & Wright, 2019, p. 118)
[10] Michael Arbib indica que parten de la doctrina de la neurona de Ramón y Cajal y la sinapsis de Sherrignton. (1987, p. 4)

el resto de la neurona. Entre la recepción de un impulso y la propagación de otro hay un retraso sináptico (McCulloch & Pitts, 1943/1990). Respecto a la excitación es interesante señalar dos cosas. En primer lugar, que los autores plantean un debate en torno a su causa, pero no la consideran relevante para el cálculo. Esto nos da la pauta de que sólo toman algunas características del modelo de la neurofisiología para crear el formalismo. En segundo lugar, que la característica que aporta la excitación es la direccionalidad del impulso nervioso que se mueve siempre en una misma dirección espacial y temporal.[11]

A continuación presentamos un esquema que ilustra las partes y relaciones de las redes neuronales "biológicas" que los autores retoman:

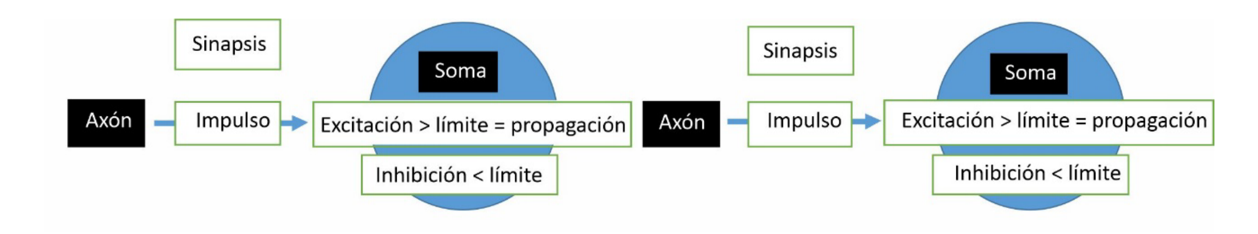


Figura 1.

Esta presentación de algunos conceptos neurofisiológicos, es retomada en la segunda parte del texto relativa a las redes sin círculos. Allí hacen una presentación, quizás a modo de sistematización, de los "supuestos físicos" para el cálculo:[12]

1. La actividad de la neurona es un proceso todo-o-nada.
2. Un cierto número fijo de sinapsis deben ser excitadas dentro de un período de adición latente para excitar a una neurona en cualquier momento, y su número es independiente de la actividad previa y la posición de la neurona.
3. El único retraso significativo en el sistema nervioso es el retraso sináptico.
4. La actividad de una sinapsis inhibitoria previene absolutamente la excitación de una neurona en ese instante.
5. La estructura de la red no cambia con el tiempo. (McCulloch & Pitts, 1943/1990, p. 101)

En lo que respecta al cálculo, el punto (1) será el puntapié de la inspiración. Que las neuronas se comporten de dicha manera posibilita la equivalencia con una lógica bivalente (con 0 y 1 como valores de verdad) en la que se toma como válido el Principio del Tercero Excluido. En relación al punto (2), lo importante es que nos indica que hay una "sumatoria" de las sinapsis previas cuyo resultado será el estímulo que recibirá el soma de otra. Si pasa su límite —independiente del estímulo[13]— el impulso es excitatorio. El punto (3) refiere a la dimensión temporal de las sinapsis. El ítem (4) nos habla del caso en el que el impulso no excede el umbral. Finalmente, (5) refiere a propiedades estructurales de la red. Según este

[11] Esto en el formalismo será expresado por el teorema 6.
[12] Según Schlatter y Aizawa (2008), la primera sección funciona como una introducción más general a la neurofisiología, lo que se sostiene en la trayectoria e investigación de McCulloch (el uso de la lógica booleana para pensar la relación entre neurona y las investigaciones de neurotransmisión en macacos, junto a Barenne, lo que les llevó a "descubrir" el fenómeno de extinción). La segunda sección lista las asunciones preliminares neurofisiológicas. La única diferencia que encuentra entre la segunda y la primera está en las velocidades de la propagación del potencial de acción. (2008, pp. 242-243)
[13] El rol excitatorio de la neurona postsináptica es independiente de la intensidad del estímulo total ya que, en cierto sentido, "rectifica" los estímulos de entrada.

ítem pareciera que no se admiten cambios en las estructuras de las redes luego de las sinapsis que se han producido.

Este momento del artículo que acabamos de exponer relativo a los supuestos físicos construye el *target* del modelo a partir de la idealización y abstracción de algunos conceptos del discurso de la neurofisiología sobre el mecanismo de comunicación neuronal. En este sentido, y también bajo el supuesto sobre el que hablamos anteriormente acerca de la identificación entre neuronas biológicas y psicones, podemos decir que el *target* del modelo es la mente-cerebro.

5. El formalismo de McCulloch y Pitts

En este apartado damos cuenta de estos objetos conceptuales y su proceso de comunicación en el marco de una descripción general del lenguaje lógico de McCulloch y Pitts. El desarrollo de este lenguaje, nos permitirá, en apartados posteriores, reflexionar acerca de la relación del discurso de la neurofisiología teórica con el de la lógica-matemática y sus efectos en la construcción del *target* del modelo.

Nos centraremos en la presentación del lenguaje que aparece en el marco del estudio de las redes sin círculos [*nets without circles*][14] y presentaremos algunos teoremas. Las redes con círculos tendrían la complejidad de remitirse a tiempos remotamente pasados y requerirían de una cuantificación temporal (Piccinini, 2004, p. 197), mientras que las redes sin círculos se limitan a expresar el tiempo a través de los predicados y los funtores.

Luego de los supuestos físicos, los autores desarrollan un formalismo. La escritura de McCulloch y Pitts es un poco compleja y, según nos indica Schlatter y Aizawa (2008), tiene errores. Haremos una interpretación de su exposición para comprender algunas cuestiones de su lenguaje. A continuación listamos sus componentes:

- ***S*** Funtor[15] definido por $\mathbf{S}(\mathbf{P})(t)\,.\!\equiv\!.\;(\mathbf{P}(\mathbf{Kx}) \wedge t = x')$.[16] "Cuyo valor para una propiedad *P* es la propiedad que vale para un número cuando *P* vale para sus predecesores." (McCulloch & Pitts, 1943/1990, p. 102)
- **Pr** es la expresión-predicado que se usa para hablar de P sin la aclaración de los paréntesis.

Como pasaba con el operador **E** en los trabajos tempranos de Pitts, si elevamos el Funtor *S* a un número, ese número nos indica la cantidad de veces que se aplicó el Funtor S: $S^2PR = S(S(PR))$ donde *PR* remite a una *expresión-predicado.*

- ***N*** nombre de la red ($\mathfrak{N}$ en el original).
- $c_1, c_2, c_3, \dots, c_n$ nombre de las neuronas
- N_i denota la propiedad de un número que una neurona c_i dispara a un tiempo que es ese número de retrasos sinápticos desde el origen del tiempo. (N_i^8 es la acción de c_i 8 instantes de tiempo desde el comienzo).

---

[14] Una idea semejante había aparecido en Turing (1936) para hablar de las máquinas, los términos que aparecen allí son "máquina circular" [*circular machine*] y "máquina libre de círculos" [*circle-free machine*] (p. 233).

[15] Este término lo toman de Carnap (1937/2007): "A los fines de expresar propiedades o relaciones de posición por medio de números, usaremos funtores" (p. 14). Actualmente, la noción de funtor es usada en teoría de categorías para hablar de morfismos entre categorías (ver Lawvere & Schanuel, 1997).

[16] La escritura original es: $\mathbf{S}(\mathbf{P})(t)\,.\equiv.\,\mathbf{P}(\mathbf{Kx})\,.\;t = x')$. Los puntos en los costados del símbolo de la equivalencia son la manera en la que se representaba la equivalencia. Actualmente se expresa sin los puntos (≡). A su vez, hay un error en el uso del paréntesis.

- $N_i(t)$ es la función que indica que c_i se dispara en un tiempo t
- N la clase de todos los $N_1, N_2, \ldots$
- $\mathbf{N_1}, \ldots, \mathbf{N_P}$ *y* $\mathbf{N_{p+1}}, \ldots, \mathbf{N_n}$, para diferenciar la acción de los aferentes periféricos (aquellas N_p que no tiene axones de entrada) de las acciones del resto de las neuronas (N_{p+1})

Dado el lenguaje anterior, la solución de la red ***N,*** los retrasos sinápticos de la misma, estará dada por oraciones S de la forma:

$$S_i\colon N_{p+1}(z_1). \equiv . Pr_i(N_1, N_2, \ldots, N_P, z_1)$$ [17]

Lo cual refiere al problema planteado por los autores como "Encontrar un método efectivo para obtener un conjunto **S** calculable o computable que constituya una solución para cualquier red dada". Y de modo converso, dado un **Pr**, se puede encontrar su realizabilidad en la red. Esto refiere al segundo problema: "Caracterizar la clase realizable **S** de una manera efectiva" (McCulloch & Pitts, 1943/1990, p. 103). Aquí puede observarse la inspiración que Turing había producido en los autores del artículo de 1943, ya que la noción de función efectivamente computable[18] surge de los trabajos de aquél.

Luego de plantear estos problemas, nos dan la definición de expresión proposicional temporal (TPE). Esta noción se vuelve importante porque es la que nos indica el tipo de proposiciones legítimas dentro del cálculo, aquellas que nos dictan el modo en que las "neuronas lógicas" se comunican. Una TPE es definida por la designación de una función proposicional temporal (TPF) de manera recursiva:

1. Una $^1p^1\,[z_1]$ es una TPE, donde p_1 es una variable de predicados.[19]
2. si S_1 y S_2 son TPE, entonces también lo son SS_1, $S_1 \vee S_2$, $S_1 \wedge S_2$ *y* $S_i \wedge\!\sim S_2$
3. Nada más es una TPE. (McCulloch & Pitts, 1943/1990, p. 103)

Como se observa en el ítem 2., estas TPE se forman a partir de las conectivas lógicas conjunción, disyunción y negación.

Luego de esta definición, los autores presentan una serie de teoremas relativos a estas TPE para poder extender su comportamiento de una manera sistémica. Por ejemplo, en el teorema 2 demuestran que las TPE son realizables en una red de orden cero. Primero se prueba para algunas TPE y luego, mediante inducción matemática, se concluye que todas las TPE son realizables.

Este lenguaje tendrá un correlato gráfico semejante a las compuertas lógicas AND, OR y AND-NOT, en cuanto modelo de implementación de funciones proposicionales binarias del álgebra booleana:

---

[17] Aquí *S* refiere a la solución de la red, no al funtor definido anteriormente, mientras que z_1 es una variable libre.

[18] Las funciones Turing-computables (es decir, computables por una máquina de Turing son funciones efectivamente computables. Lo que no es obvio es la conversa y que constituye la famosa Tesis de Turing: si una función es efectivamente computable, entonces es Turing-computable. Boolos, Burgess y Jeffrey definen a una función *f* efectivamente computable como aquella que puede dar una lista de instrucciones para poder determinar el valor de *f*(*n*) para cada argumento de *n* (ver "Turing Computability" en Boolos, Burgess, & Jeffrey, 2007, pp. 3-33).

[19] z_1 según definieron unas páginas antes en el texto es una variable libre y $^1p^1$ es un predicado dentro del conjunto mayor *PR*.

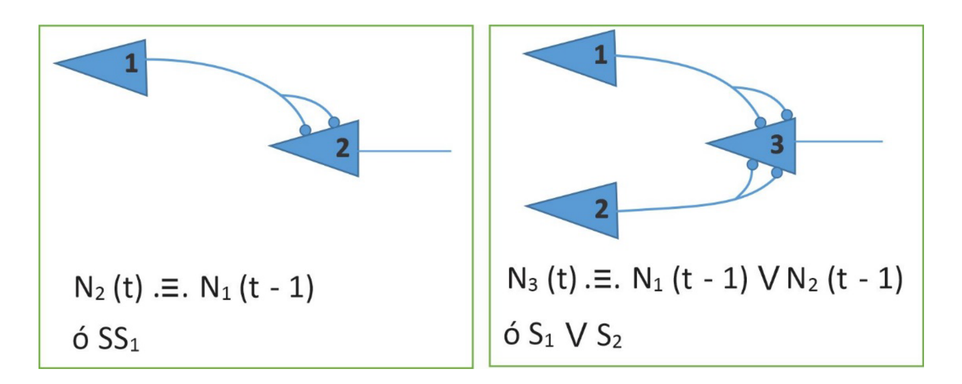


Figura 2.

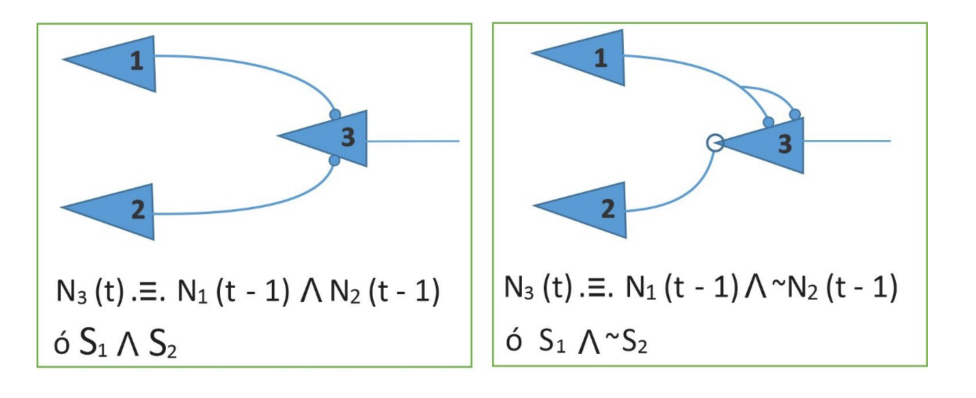


Figura 3.

La escritura en la que se usa la función $N_i(t)$ nos permite indicar el momento en el que se dispara la neurona en relación a un tiempo anterior a *t*. De modo que esta escritura da cuenta de sus posibles procesos anteriores como equivalentes a la función del disparo en un tiempo *t*. Mientras que el funtor S nos permite resumir esa expresión temporal. Este nos indica el disparo de la neurona haciendo referencia a las sucesiones de tiempos previas. Como se anticipó, si S no está elevando a ningún número, significa que hubo sólo una sucesión en el tiempo. Si está elevado a la 2, hubo dos sucesiones de tiempo, y de manera general, si está elevado a la *n*, hubo *n* sucesiones de tiempo.

Este correlato espacial expresado en las figuras 2 y 3 de las proposiciones temporales se justifica, de alguna manera, posteriormente en el teorema 6. Este nos indica que la facilitación y la suma temporal pueden ser reemplazadas por la suma espacial. Tanto la facilitación que remite a la amplitud en la recepción de señales neuronales, como la suma espacial y temporal que remiten la suma de las señales previas en cuanto a posición y tiempo, producen alteraciones en los ciclos. Sin embargo, las neuronas de McCulloch-Pitts abstraen estos cambios "factuales" y los autores conciben a sus redes invariantes respecto a estas alteraciones con motivos de simplificación representacional (McCulloch y Pitts 1943/1990, p. 101).

Como observación general, decimos que McCulloch y Pitts utilizan un lenguaje proposicional semejante al álgebra booleana para indicar la manera en que las neuronas lógicas se comunican, y un lenguaje de predicados para dar cuenta del comportamiento temporal de las neuronas. El modo de exposición de las ideas es semejante a la presentación de un cálculo lógico: detallan los componentes de su lenguaje, dan definiciones recursivas y demuestran teoremas. Sin embargo para algunos autores no se trataría de un cálculo en sentido estricto ya que un "sistema deductivo está destinado a capturar inferencias lógicas entre la clase de expresiones permitidas por la gramática del cálculo" (Piccinini, 2004,

p. 189). Si bien se demuestran teoremas no hay un conjunto de reglas sintácticas o reglas de inferencias a partir de las cuales se lleven a cabo las demostraciones. Otra de las críticas que se ha hecho al uso de la lógica apareció en una reseña temprana del artículo escrita por Frederic Fitch (1944):

> La lógica simbólica se usa con libertad e, incluso, con bastante descuido. La notación es innecesariamente complicada y el número de errores tipográficos es tan grande que el revisor no pudo descifrar varias partes cruciales del artículo. En algunos lugares los errores parecen ser de naturaleza lógica en lugar de meramente tipográficos, por ejemplo, la derivación de la fórmula (3) en la página 125. En cualquier caso, no existe una construcción rigurosa de un "cálculo lógico". (p. 49)

Sin embargo, a pesar de estas críticas, el lenguaje ha sido usado como un puntapié para desarrollar lenguajes de diseño lógico y redes neuronales, como veremos en el siguiente apartado.

6. Los efectos del modelo

Para Piccinini (2004) uno de los aportes del modelo de McCulloch y Pitts ha sido contribuir al desarrollo de una técnica para el diseño lógico de computadoras. El trabajo de 1943 "encuentra simpatía en una audiencia de científicos interesados en la epistemología pero entrenados en matemática o ingeniería más que en neurofisiología, como Norbert Wiener y John von Neumann." (Piccinini, 2003, p. 62). Esto se expresa en las Conferencias Macy y el Simposio Hyxon (1948), de las cuales surge la idea de que el estudio del cerebro y de las computadoras pertenece a la misma ciencia (Piccinini, 2003) y de esta manera se posibilita el desarrollo de distintos tipos de estudios:

1. La teoría de los autómatas de von Neumann.
2. El diseño de redes específicas usando las neuronas idealizadas para
3. funciones mentales específicas (ej.: modelo de la visión de McCulloch y Pitts de 1947 y el Perceptron de Rosenblatt de 1958[20]. Esto dio origen al desarrollo de estudios de redes neuronales.
4. El desarrollo de programas de computadora que ejecutan funciones > mentales
5. La neurofisiología experimental.[21] (Piccinini, 2003, p.110)

Los debates planteados por el artículo de 1943 estuvieron muy presentes en las Conferencias Macy en las que se llevaron a cabo importantes desarrollos teóricos para la investigación en IA. Es por ello que puede considerarse que "el modelo McCulloch-Pitts no engendra una notable posteridad ni en la lógica ni en la neurofisiología pero sí en la inteligencia artificial" (Dupuy, 2000, p. 62). En este sentido, podría decirse que si el aporte que ha dejado el modelo McCulloch-Pitts ha sido principalmente en el campo del diseño lógico y la teoría de los autómatas, el alcance representativo del modelo excedió la mente-cerebro de los autómatas naturales. El modelo produjo la posibilidad, o fue uno de los

[20] Si bien el Perceptron es un tipo particular del modelo neuronal McCulloch-Pitts, Rosenblatt, su creador, se demarcó de dichos autores por considerar su enfoque insuficientemente biológico (ver Dupuy, 2000, p. 62)

[21] Quizás la influencia más directa pueda ser el caso de Jerome Lettvin, tal y como comenta Piccinini: "El principal trabajo de Lettvin fue en este campo. Con Pitts, McCulloch y Humberto Maturana, Lettvin interpretó las señales que viajaban a través del nervio óptico de una rana como si fuera una implementación de los mecanismos descritos en el *Pandemonium* de Selfridge (Lettvin, Maturana, McCulloch y Pitts 1959). Su trabajo fue un gran avance en la interpretación de las propiedades de respuesta de las neuronas y se convirtió en un modelo para muchos trabajos en neurofisiología durante las décadas posteriores." (Piccinini, 2003, p. 111)

aportes claves en el diseño de autómatas artificiales, lo que muestra que un modelo puede producir un fenómeno *target* no esperado por los autores del mismo.

7. La aparición de las neuronas lógicas

Si bien consideramos que el *target* del modelo McCulloch Pitts es el fenómeno de la mente-cerebro, aparecen varias entidades en juego en dicha propuesta. Si hacemos un repaso, en el artículo de 1943 encontramos cuatro tipos de entidades conceptuales: las neuronas biológicas y sus redes, los psicones como átomos mentales y soporte de los procesos mentales, las proposiciones y, por último, las *neuronas lógicas o formales* y sus redes, producto de la equivalencia formal.

Donald Perkel (1988), en un artículo que repasa el legado teórico de McCulloch, llama a las neuronas McCulloch-Pitts "neuronas lógicas" para remarcar el aspecto formal de las mismas. En este sentido algunos autores han hablado de "neuronas formales" (Dupuy, 2000) o redes neuronales formales (von Neumann, 1951). Los propios creadores del modelo hablaron del carácter formal del mismo cuando distinguieron entre "equivalencia factual" y "equivalencia formal", distinción que tematizaremos en el próximo apartado.

También apelaron a la idea de "redes ficticias" para hablar de la substitución de redes con alteraciones por otras que no las tuvieran: "para redes que experimentan ambas alteraciones, podemos sustituir redes ficticias equivalentes compuestas de neuronas cuyas conexiones y límites son alterados" (McCulloch & Pitts, 1943/1990, p. 101). Con "alteraciones" se referían a aquellas cuestiones que modificaban toda la red: por un lado, a la facilitación y la extinción de las neuronas en las que la información temporalmente antecedente afecta la reacción posterior; y por otro lado, el aprendizaje que posibilita que un estímulo que anteriormente fue inadecuado deje de serlo. Con ficticias entonces se refieren en este caso a la idealización de las redes para simplificar su mecanismo. Tal como rastrea Dupuy (2000) en el recorrido teórico de McCulloch: si apelaba a la idea de "ficticias" era para referirse a las entidades idealizadas. Sin embargo, de acuerdo a la revisión que hace McCulloch del *paper* de 1943 a mediados de los años '60, la idealización de las neuronas no significaba que no se tratara de neuronas físicas. En este contexto el neurofisiólogo reniega de cierto devenir "desencarnado" e "incorpóreo" de la recepción de su *paper* en la teoría de los autómatas, que se presenta como un contrapunto con las posiciones de von Neumann y el uso que hace del lenguaje McCulloch-Pitts para el diseño de los primeros autómatas artificiales. (pp. 59, 61). Entonces el hecho de que los autores consideren a las redes como ficticias puede entenderse como una postura ficcionalista que considera a un modelo, en tanto idealización, como una ficción, en un marco que considera a la actividad misma de producir modelos como ficcional. O también puede entenderse que "ficticias" refiere a que tratan con entidades ficcionales: las neuronas lógicas o formales.

Sistema nervioso como conjunto de neuronas (*neuronas biológicas* y sus redes neuronales).	*Psicones* como átomos mentales. Las neuronas como soporte de la mente y por ende de los procesos mentales como el aprendizaje.	*Proposiciones* de la lógica proposicional bivalente.	Equivalencia formal: *neuronas lógicas* y sus redes.

Tabla 1: Tipo de entidades involucradas en el modelo de Mculloch y Pitts (1943/1990)

Entonces, si nos centramos en la recepción del lenguaje de McCulloch y Pitts como posibilitador del desarrollo de la teoría de los autómatas, podemos pensar lo siguiente: que, o bien las neuronas lógicas surgieron como representación de la mente-cerebro de los autómatas naturales y que, a su vez, posibilitaron la existencia de otro *target*, los autómatas artificiales; o bien, que los autómatas artificiales también son un modelo de la mente-cerebro humana. La respuesta a esta pregunta dependerá de nuestro vínculo con el humanismo y la concepción de la relación humano-máquina, lo cual no es debate de este artículo.

Lo que sí está claro es que los desarrollos teóricos de McCulloch y Pitts contribuyeron con gran protagonismo al despliegue de las investigaciones en inteligencia artificial. En este sentido, el fenómeno *target* del modelo parecería ser las neuronas lógicas. Y es en este sentido, que las neuronas lógicas en tanto objetos conceptuales no son triviales: implicaron un proceso de construcción y el desarrollo de nuevas áreas de investigación.

8. La equivalencia formal

La equivalencia formal es una de las cuestiones centrales del texto. Aparece incluso como aspecto clave en el *abstract* mismo, en donde se plantea como un objetivo del trabajo mostrar que

> Muchas opciones particulares entre los posibles supuestos neurofisiológicos son equivalentes, en el sentido de que por cada red que se comporta bajo una suposición, existe otra red que se comporta de acuerdo con la otra y da los mismos resultados, aunque quizás no al mismo tiempo. (McCulloch & Pitts, 1943/1990, p. 99)

Ya en el texto, la cuestión de la equivalencia formal es introducida luego de la presentación de dos posibles explicaciones acerca de la inhibición. Al respecto nos dicen que: "Dado que nos interesan las propiedades de las redes que son invariantes bajo equivalencia, podemos hacer las suposiciones físicas que sean más convenientes para el cálculo" (McCulloch & Pitts, 1943/1990, p. 100). Es decir, que si bien las suposiciones físicas se toman como punto de inspiración para el cálculo, no hay una pretensión de que coincidan exactamente sino que se pueden modificar en función de las necesidades de aquel, lo que explicamos en el apartado anterior cuando nos referimos al carácter ficticio de las redes.

McCulloch y Pitts distinguen la equivalencia formal de la explicación o equivalencia factual. En este contexto, aparece cierta ambigüedad en sus ideas, sin embargo, todas apuntan a una noción de identidad por correspondencia o identificación que, aunque no sea puesta explícitamente en términos de funciones biyectivas o isomorfismos, nos recuerda mucho a dichas ideas. ¿Qué significa la *equivalencia formal*? Al respecto aparecen tres ideas que nos parece que ayudan a vislumbrar la cuestión:

> 1. Las relaciones psicológicas que existen entre la actividad nerviosa *corresponden* a la relación entre proposiciones. La utilidad de la representación dependerá de la *identidad* de esas relaciones con la lógica de proposiciones;
> 2. Para cada reacción de una neurona hay una aserción *correspondiente* de una proposición -e implicará otra proposición-.
> 3. Las alteraciones como la facilitación, extinción y aprendizaje no afectan a las conclusiones que se saquen del *tratamiento formal* de la actividad de las redes nerviosas. (McCulloch & Pitts, 1943/1990, pp. 100–101)

Si tenemos en cuenta lo anteriormente dicho, aparecen varios sentidos respecto de la equivalencia formal, que se asocia a las nociones de identidad y correspondencia. En primer lugar, se habla de la correspondencia entre las proposiciones y las relaciones psicológicas que suponen en la actividad nerviosa. En segundo lugar, la correspondencia o identidad se da entre la lógica de proposiciones y las reacciones de las neuronas. Y, en tercer lugar, se remarca la relevancia del tratamiento formal y se aclara que las alteraciones (facilitación, extinción y aprendizaje) que implicarían un cambio en la información recibida por las neuronas y que afectarían a la red en su totalidad, no son tenidas en cuenta. Esto se debió a la necesidad de simplificar el cálculo. A su vez, se dice que la utilidad de la representación dependerá de la identidad entre las relaciones psicológicas o la actividad nerviosa de las neuronas y las proposiciones. De esta manera, la equivalencia formal es la garantía para que el formalismo sea una representación de la actividad neuronal y, en consecuencia, de los procesos psicológicos humanos soportados en los átomos mentales o psicones. De este modo se constituye en la explicación científica que garantiza el mecanismo de la mente-cerebro pero al mismo tiempo, en un criterio de isomorfismo del modelo, es decir, la correspondencia uno-a-uno entre el modelo y el *target* (Potochnik, Colombo, & Wright, 2019, p. 117).

A su vez, se dice que la utilidad de la representación dependerá de la identidad entre: relaciones psicológicas y la actividad nerviosa con las proposiciones, lo cual puede representarse según el siguiente esquema:

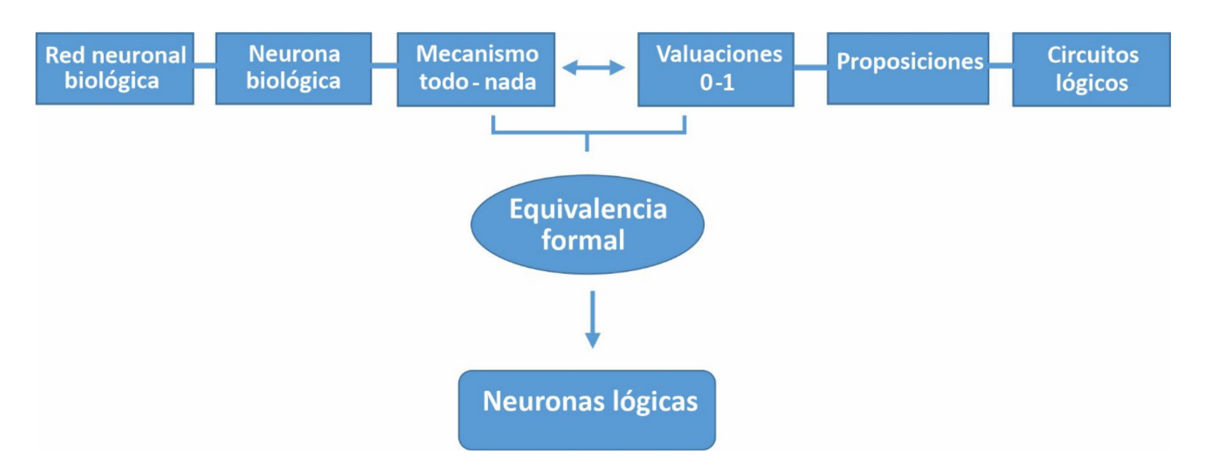


Figura 4.

De esta manera, la equivalencia formal puede entenderse ya no como el criterio de isomorfismo según el cual las neuronas lógicas representarían el *target* dado por la mente-cerebro, sino como el isomorfismo entre neuronas biológicas, psicones y proposiciones que producen una nueva entidad conceptual: las neuronas lógicas. La correspondencia entre las neuronas biológicas y el comportamiento de la valuación de las proposiciones lógicas, daría lugar a un *entre* que toma aspectos del comportamiento de ambas partes. Hemos visto que en este marco la equivalencia formal ha cumplido un rol fundamental: en tanto explicación científica, en tanto criterio de isomorfismo entre el modelo y el fenómeno target y finalmente como criterio de identidad de las entidades conceptuales producidas en el marco del modelo: las neuronas lógicas.

9. Conclusión

En este trabajo hemos visto que el desarrollo del modelo neuronal de McCulloch y Pitts ha producido varios aportes los cuales retomamos, desarrollamos y, a partir de ellos, analizamos las implicancias de dicho modelo en la historia de la ciencia. En primer lugar, en tanto teoría computacional de la mente hemos visto que esta se debió a la aplicación de la noción de computación de Turing. Esta teoría de la mente se trató de una explicación mecanicista que se insertó en los desarrollos de la neurofisiología teórica de Ramón y Cajal, Barenne, entre otros. La forma en que se llevó a cabo dicha teoría fue a partir de la resolución del problema mente-cuerpo asociando una teoría de la mente que venía desarrollando McCulloch —la teoría de los psicones— a la teoría de la sinapsis neuronal. A su vez estas unidades nerviosas-mentales fueron asociadas a proposiciones lógicas a partir de la correspondencia establecida entre el mecanismo todo-nada de las neuronas y la lógica bivalente del álgebra de Boole. A su vez, influenciados por los estudios de biología matemática de Woodger, Rashevsky y Householder desarrollaron modelos de redes neuronales a través de la producción de un lenguaje lógico, para el cual fueron también influencias importantes Russell y Carnap.

Por otro lado, el uso del lenguaje II de Carnap y la lectura de los *Principia Mathemathica* posibilitaron la creación de una escritura para indicar información temporal al *interior* de las proposiciones. O, dicho de otra manera, para incluir en el *contenido* de las proposiciones, la dimensión temporal de la comunicación de las neuronas lógicas. A cada fórmula con su operación, le correspondía una imagen del tipo presentado en el apartado anterior. A su vez, la influencia de Russell-Whitehead y Carnap permitió organizar el texto en forma de cálculo con definiciones recursivas, deducción de teoremas y el uso de la inducción matemática como forma de generalizar. Sin embargo esta lectura tiene un contrapunto en la crítica de Fitch y Piccinini acerca de la necesidad de reglas sintácticas para considerar a un lenguaje un cálculo lógico.

La confluencia de distintos leguajes lógico-matemáticos con algunas ideas de la neurofisiología produjo un nuevo lenguaje que posibilitó la existencia de un nuevo tipo de neuronas. En particular, la equivalencia formal del mecanismo todo-nada de la neurona biológica con la lógica binaria (0-1) del álgebra de Boole. Pero no sólo la correspondencia en el mecanismo de una lógica bivalente posibilitó la construcción de un nuevo lenguaje. Pareciera que también la semejanza en la representación visual de las neuronas biológicas y los diagramas de flujo de las compuertas OR, AND y AND-NOT. La *yuxtaposición* de la representación gráfica de la neurona biológica con los diagramas posibilitó una escritura espacial que sintetizó dos representaciones en una nueva. Estos nuevos objetos no eran exactamente iguales a las neuronas biológicas ni exactamente iguales a la representación gráfica de las compuertas lógicas de Boole.

En este sentido, la equivalencia formal, que tuvo inicialmente la motivación de explicar los procesos mentales, produjo una nueva entidad conceptual: las neuronas lógicas, que a su vez posibilitó el desarrollo de nuevas investigaciones científicas. Es así que posibilitaron en la historia de la cibernética el paso del uso de la noción de computación para entender la mente al uso de la noción de mente para diseñar mecanismos computacionales.

En tanto modelo científico que pretende representar el mecanismo de la mente-cerebro, podría pensarse que el lenguaje de McCulloch y Pitts lleva a cabo dicha representación. Pero también podría pensarse que, en la creación del *target* del modelo, se

amplía el alcance representativo de éste último. Ya no sólo representa la mente-cerebro sino también la posibilidad del desarrollo de máquinas inteligentes. Para ello fue necesario que la equivalencia formal funcione como explicación científica en contraposición a la explicación factual y que haya una concepción mecanicista de dicha equivalencia según la cual mente, cerebro y mecanismo convergen. Podemos pensar, entonces, que lo que ahora es recordado como el modelo neuronal McCulloch-Pitts es producto de conjugar diferentes disciplinas de tradiciones formales y experimentales.

Referencias

Abraham, T. H. (2002). (Physio)Logical Circuits: The intellectual origins of the McCulloch-Pitts Neural Networks. *Journal of the History of the Behavioral Sciences*, 38 (1), 3–25. https://doi.org/10.1002/jhbs.1094

Arbib, M. (1987). *Brains, machines and mathematics*. Springer Verlag.

Barberis, S. (en prensa). La explicación mecanicista. En prensa.

Branca, M. I., Ramírez, A. O., y Vilatta, M. E. (2015). Modelos de explicación en psicología cognitiva y neurociencias. *Anuario de Investigaciones de la Facultad de Psicología*, 2 (1), 176-191. https://www.revistas.unc.edu.ar/index.php/aifp

Boolos, G., Burgess, J., Jeffrey, R. (2007). *Computability and Logic*. Cambridge University Press.

Carnap, R. ([1937] 2007). *Logical Syntax of Language*. Routledge.

Cassini, A. (2018). Models without a Target. *Artefactos. Revista de estudios de la ciencia y la tecnología*. 7 (2), 185-209. http://dx.doi.org/10.14201/art201872185209

Craver, C. (2007). *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*. Oxford, Clarendon Press.

Dupuy, J-P. (2000). *The mechanization of the mind: On the origins of cognitive science*. Princeton University Press.

Fitch, F. (1944). Review of McCulloch and Pitts 1943. *Journal of Symbolic Logic*. 9 (2), 49–50. https://doi.org/10.2307/2268029

Householder, A. (1941a). A theory of steady-state activity in nerve-fiber networks: I. Definitions and preliminary lemmas. *Bulletin of Mathematical Biophysics*, 3, 63–69. https://doi.org/10.1007/BF02478220

Householder, A. (1941b). A theory of steady-state activity in nerve-fiber networks II: The simple circuit. *Bulletin of Mathematical Biophysics*, 3, 105–112. https://doi.org/10.1007/BF02478168

Householder, A. (1941c). A theory of steady-state activity in nerve-fiber networks III: The simple circuit in complete activity. *Bulletin of Mathematical Biophysics*, 3, 137–140. https://doi.org/10.1007/BF02477933

Householder, A. (1942). A theory of steady-state activity in nerve-fiber networks IV: N circuits with a common synapse. *Bulletin of Mathematical Biophysics*, 4, 7–14. https://doi.org/10.1007/BF02477350

Kleene, S. C. (1956), Representation of Events in Nerve Nets and Finite Automata en C. E. Shannon y J. McCarthy (Eds.), *Automata Studies*, (3-42). Princeton University Press.

Lawvere, F. W. & Schanuel, S. (1997). *Conceptual Mathematics: A First Introduction to Categories.* Cambridge University Press.

(McCulloch & Pitts, 1943) McCulloch, W., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, *5*(4), 115-133. https://doi.org/10/djsbj6

McCulloch, W. y Pitts, W. (1990). A logical calculus of de ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 52 (l/2), 99-115. (Obra original de 1943) https://doi.org/10.1007/BF02459570

Northrop, F. S. C. (1931). *Science and first principles.* Macmillan.

Perkel, D. (1988). Logical neurons: the enigmatic legacy of Warren McCulloch. *Trends in Neurosciences*, 11 (1), 9-12. https://doi.org/10.1016/0166-2236(88)90041-0

Piccinini, G. (2003). *Computations and computers in the sciences of mind and brain.* [Tesis de doctorado, Universidad de Pittsburgh]. https://www.researchgate.net/publication/301340375_Computation_and_Computers_in_the_Sciences_of_Mind_and_Brain

Piccinini, G. (2004). The first computational theory of mind and brain: a close look at McCulloch and Pitts' "Logical calculus of de ideas immanent in nervous activity". *Synthese*, 141, 175–215. https://doi.org/10.1023/B:SYNT.0000043018.52445.3e

Potochnik, A., Colombo, M., Wright, C. (2019). *Recipes for Science. An Introduction to Scientific Methods and Reasoning*. Editorial Routledge.

Schlatter, M. y Aizawa, K. (2008). Walter Pitts and "A Logical Calculus". *Synthese*, 162, 235–250. https://doi.org/10.1007/s11229-007-9182-9

von Neumann, J. (1945). *First Draft of a Report on the EDVAC*. Technical Report, Moore School of Electrical Engineering, University of Pennsylvania. En línea: http://abelgo.cn/cs101/papers/Neumann.pdf

von Neumann, J. (1951). The General and Logical Theory of Automata. En L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior* (pp. 1–41). Wiley.

Turing, A. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, *42*(1), 230-265. En línea: https://www.cs.virginia.edu/~robins/Turing_Paper_1936.pdf

Zamora-Cárdenas, W., Zumbado, M., Trejos-Zelaya (2020). McCulloch-Pitts Artificial Neuron and Rosenblatt's Perceptron: An abstract specification in Z. *Revista Technology inside by CPIC*, 5 (5), 16-29. https://cpic-sistemas.or.cr/revista/index.php/technology-inside/article/view/36/27

Declaraciones

Conflictos de interés: La autora declara que no existen conflictos de interés.

La aguja de Hans Christian Ørsted

Traducción de sus dos primeros artículos sobre electromagnetismo

Rodrigo Andrés Bautista Rincón[1]

Recibido: 5 de julio de 2022
Aceptado: 1 de noviembre de 2022

Resumen: El presente escrito da un panorama acerca de la electricidad y el magnetismo al terminar el siglo XVIII y al iniciar el siglo XIX, contexto en el cual, se tiene la invención de la pila de Alessandro Volta, y el desarrollo de la experiencia de la aguja de Hans Christian Ørsted, ambos hechos fundamentales en el encausamiento de las prácticas experimentales y en la constitución de los cuerpos teóricos de las ciencias. De esta manera, se presentan los dos textos escritos por el científico danés en el año de 1820 (uno traducido del latín al inglés), los cuales fueron publicados por la revista *Annals of Philosophy*. En ellos se detalla en primer lugar la experiencia que tuvo lugar y algunas de las ideas que estableció Ørsted para darle un sentido al fenómeno observado. Posteriormente, se establecen nuevos experimentos en el marco de lo que asumió como electromagnetismo.

Palabras clave: electricidad, magnetismo, electromagnetismo, Hans Christian Ørsted.

Title: Hans Christian Ørsted's Needle: Translation of his first two articles on electromagnetism.

Abstract: The present writing gives an overview of electricity and magnetism at the end of the 18th century and at the beginning of the 19th century, a context in which there is the invention of the battery by Alessandro Volta, and the development of the experience of the needle by Hans Christian Ørsted, both fundamental facts in the channeling of experimental practices and in the constitution of the theoretical bodies of the sciences. In this way, the two texts written by the danish scientist in the year 1820 (one translated from Latin to English), which were published by the journal Annals of Philosophy, are presented. They first detail the experience that took place and some of the ideas that Ørsted established to give meaning to the observed phenomenon. Subsequently, new experiments are established within the framework of what he assumed as electromagnetism.

Keywords: electricity, magnetism, electromagnetism, Hans Christian Ørsted.

1. Introducción

La expansión comercial de la segunda parte del siglo XIX supuso la búsqueda de un descenso en los precios de transporte y una rápida distribución que llevaran a una mayor acumulación de capital. Para ello, era fundamental la invención de un motor que dependiera de una nueva fuente de energía diferente al trabajo suministrado por un gas en un motor a vapor. De tal forma, el trabajo científico estuvo abocado al desarrollo de experimentos con

---

[1] Universidad Pedagógica Nacional. Bogotá, Colombia.
rabautistar@gmail.com | 0000-0001-5845-9349
Bautista Rincón, R. A. (2022). La aguja de Hans Christian Ørsted: Traducción de sus dos primeros artículos sobre electromagnetismo. *Epistemología e Historia de la Ciencia, 7*(1), 41–51. https://revistas.unc.edu.ar/index.php/afjor/article/view/38222

electricidad y magnetismo como posibles fuentes de energía, sin embargo, existía confusión en cuanto a términos, se tenía una gran dificultad para elaborar montajes o replicar los resultados que otros tuvieron, además, los fenómenos eléctricos y los fenómenos magnéticos parecían indiferenciados porque sus manifestaciones eran similares.

Para esta época, ya se hablaba de cuerpos cargados, pero muy a pesar de ello, no se tenía conocimiento de la naturaleza de la carga, por tanto, se consideró como un ente definido que podría poseer carácter de fluido, que asumía dos formas las cuales se designaron por arbitrariedad como carga positiva y carga negativa. El exceso de alguna de las dos tipologías implicaba su presencia en los cuerpos.

En este camino, los fenómenos de atracción y repulsión de cuerpos cargados, fueron abordados bajo estas formulaciones y formalizados con la teoría de acción a distancia de Newton.

Posteriormente, los científicos observaron que al conectar dos cuerpos que poseían diferente carga a través de un alambre, se tenía el movimiento de un fluido en este. Tal hecho desencadenó el siguiente cuestionamiento: ¿son dos fluidos o solamente uno? Esto supuso entender y diferenciar lo estático y dinámico dentro de la electricidad, y acrecentó los problemas en cuanto a poder explicar la naturaleza de aquello que se hacía presente en la materia (Berkson, 1981).

De manera similar, para el magnetismo también se consideraban dos fluidos, uno que se le denominó norte y otro sur (Millikan, 1944). El problema de esta distinción, era que en la práctica el norte y el sur no se podían aislar para ser analizados, dado que no se podían separar, lo que dificultaba la comprensión de estos fluidos que producían fenómenos extraordinarios y que solían ser confundidos con los de la electricidad.

Los escritos de Faraday hacia los años 1821 y 1822, dejan entrever esta caótica situación en el entendimiento de los términos y sus efectos:

> Aquellos que consideran la electricidad como un fluido, o como dos fluidos, piensan que una corriente o corrientes de electricidad están pasando a través del alambre durante todo el tiempo que forma la conexión entre los polos de un aparato activo. Hay muchos argumentos a favor de la materialidad de la electricidad, y pocos contra ella; pero todavía es sólo una suposición; y será bueno recordar, mientras se desarrolla el tema del electro-magnetismo, que no tenemos prueba de la materialidad de la electricidad, o la existencia de cualquier corriente a través del alambre.[2].(Faraday, 1821, p. 196)

En este panorama tan ambiguo, los científicos también se interesaron por la relación entre la electricidad y el magnetismo, pues se tenía evidencia de que los rayos producían efectos magnéticos, esto a partir de experiencias de personas como Benjamin Franklin, quien magnetizó agujas de costura al descargar una botella de Leyden; igualmente, se tenían experiencias con agujas magnetizadas y agujas cargadas, las cuales se orientaban en la misma dirección, y aún más, se llegó a abordar la electrólisis mediante el uso de un imán en lugar de una batería.

De esta forma, el inicio del siglo XIX dispuso como propósito para la Ciencia entender qué era la electricidad y el magnetismo, y cuál era su correspondencia, teniendo como base que, en algunos casos, un imán o una corriente generada mediante una batería producían los mismos efectos.

---

[2] Traducción propia del autor.

2. Una aguja en movimiento

En el final del siglo XVIII, la creación de la pila voltaica derivó en la realización de diversas experiencias, particularmente, en el ámbito de la electroquímica. No obstante, hacia el año de 1802, Gian Doménico Romagnosi llevaría a cabo un experimento que debería situarse como la apertura hacia el electromagnetismo, pues como lo relata Andrade (2001), detalló la desviación de una aguja magnética como consecuencia de la interacción con una corriente voltaica, observaciones que publicaría en dos escritos dispuestos en la *Gazzeta di Trento* (agosto de 1802) y la *Gazzeta di Rovereto* (octubre de 1802):

> Preparó la pila del señor Volta, construida con piezas redondas de zinc y cobre, alternando con trozos de franela húmeda con agua impregnada de sal amoniacal; un hilo de plata, dividido en varios pedazos como una cadena, se conectó a la pila. El último segmento de la cadena pasó a través de un tubo de vidrio, y desde su terminación exterior salía un broche de plata pura unido a la cadena. Una vez hecho esto, tomó una aguja magnética ordinaria, en forma de brújula náutica, encajonada en una tabla de madera cuadrada; y levantando el cristal que la encerraba, la puso sobre un aislante de vidrio, cerca de la pila antes mencionada. Luego, tomando la cadena de plata y sosteniéndola por el tubo de vidrio mencionado anteriormente, puso su extremo o perilla en la aguja magnética; y manteniendo el contacto durante unos segundos, hizo que la aguja se desviará unos pocos grados de la dirección de los polos. (Andrade, 2001, pp. 88–89)

A pesar de la importancia de las observaciones, este trabajo no iría más allá de pequeñas menciones realizadas entre 1804 y 1830.

En el año de 1820, Hans Christian Ørsted realiza una publicación en latín detallando una experiencia similar a la de Romagnosi, la cual es traducida posteriormente al inglés para ser dispuesta en los *Anales de Filosofía* en el mes de octubre. Este trabajo marcaría un antes y un después en la historia de la ciencia, pues junto con la pila de Volta, llevaría a la formulación de múltiples experiencias, observaciones y teorizaciones sobre el comportamiento y relación de la electricidad y el magnetismo.

Dadas las semejanzas de las conclusiones de Romagnosi y Ørsted (no así la naturaleza de sus experimentos, pues en el primero se tuvo un contacto y la explicación de los hechos se puede dar desde la electrostática, mientras que en el segundo no se tuvo contacto de ningún tipo entre materiales y todo se atribuiría a una acción dispersa en el espacio), se ha tenido una discusión en la comunidad académica sobre quién descubrió el electromagnetismo y por tanto, quién realmente merece el crédito. Sin embargo, asumiendo la postura de Kuhn (1962/2004), induce al error sugerir que descubrir algo se reduce a un acto único, a un individuo o a un instante temporal, porque descubrir un nuevo tipo de fenómeno es necesariamente un suceso complejo que entraña reconocer tanto que algo es, como qué es (p. 107).

Así pues, se consideran los escritos de 1820 de Hans Christian Ørsted, en los cuales detalla su experiencia y en donde trata de dar una base teórica a los hechos observados. En ellos, es posible destacar tres formulaciones. Por una parte, en un primer instante (1820) recurre a la formulación de un "conflicto eléctrico" generado en el cable, que se distribuye por el espacio de manera circular, el cual afectaba tanto a los cuerpos que se resistían a su influencia, como aquellos en donde residía el magnetismo, manifestándose en términos de una fuerza. Posteriormente, en 1821, Ørsted trata de formalizar sus consideraciones como fuerzas transversales que se generan por el tránsito de una corriente eléctrica en el

conductor, las cuales, alteran el estado de la aguja, y en este camino, traza un intento por entender el magnetismo de la tierra. Por último, menciona diferentes disposiciones experimentales en donde no solo se limita a analizar el efecto de un cable conductor sobre una aguja magnética evidenciado en movimiento y declinación, sino que trata de visualizar un efecto similar de parte de un imán sobre un cable conductor ligado a una pila.

Con los efectos observados, Ørsted concibe que lo que se denomina electricidad no era realmente eso, en sentido estricto, pues la fuerza en el circuito actuaba en una forma particular, que convenientemente denominó como magnética. Como la acción magnética se da en una línea recta, es decir, las fuerzas opuestas se mueven en direcciones opuestas, mientras que las fuerzas en la experiencia, fluyen incesantemente y forman un círculo, llevan a Ørsted a llamar a este efecto *Electro-magnetismo* (Jackson, Jelved, & Knudsen, 1998).

Cabe mencionar que la visión de Ørsted estaba ligada a la metafísica kantiana (Brain, Cohen, & Knudsen, 2007), en la que se consideraba que las fuerzas son inmateriales de tal manera que no pueden ser intuidas ni en el tiempo ni en el espacio, es decir, son inaccesibles y solamente es posible percibir sus efectos, los cuales son empíricos. En este orden de ideas, Ørsted entendía la electricidad y el magnetismo como fuerzas ocultas cuyos efectos, podían transmitirse por medio de ondulaciones (idea que se apoyaba en una visión de espacio como lleno de materia).

3. Primer escrito (octubre de 1820)

Experimentos sobre el efecto de una corriente de electricidad sobre la aguja magnética [3]

Por Hans Christian Ørsted Caballero de la Orden de Dannebrog, Profesor de Filosofía Natural, y Secretario de la Royal Society de Copenhague

Los primeros experimentos sobre el tema que en este momento quiero explicar, fueron hechos por mí el último invierno, mientras daba clases sobre electricidad, galvanismo y magnetismo, en la Universidad. Parecía demostrado por estos experimentos que la aguja magnética era movida de su posición por el aparato galvánico, pero que el círculo galvánico debía estar completo, y no abierto, último método que fue probado en vano hace algunos años por muy célebres filósofos. Pero como estos experimentos se hicieron con un aparato débil, y no eran, por lo tanto, suficientemente concluyentes, considerando la importancia del tema, me asocié con mi amigo Esmarck para repetirlos y extenderlos mediante una batería galvánica muy potente, proporcionada por nosotros. El Señor Wleugel, Caballero de la Orden de Dannebrog, estuvo presente, y colaboró en los experimentos. Estaban presentes también el Señor Hauch, un hombre muy hábil en las Ciencias Naturales, el Señor Reinhardt, Profesor de Historia Natural, el Señor Jacobsen, Profesor de Medicina, y ese químico muy hábil, el Señor Zeise, Doctor en Filosofía. A menudo había hecho experimentos por mí mismo; pero cada hecho que había observado se repitió en presencia de estos caballeros.

El aparato galvánico que empleamos consistía en 20 cubetas de cobre, la longitud y altura de cada una era de 12 pulgadas; pero la anchura apenas excedía 2 ½ pulgadas. A cada cubeta se le suministraron dos placas de cobre, tan dobladas que podían llevar una varilla de

---

[3] Traducido de un informe impreso redactado en latín por el autor, y transmitido por él al editor.

cobre, que soporta la placa de zinc en el agua de la siguiente cubeta. El agua de las cubetas contenía 1/60 de su peso en ácido sulfúrico, y una cantidad igual de ácido nítrico. La porción de cada placa de zinc sumergida en el agua es un cuadrado cuyo lado es de aproximadamente 10 pulgadas de longitud. Un aparato más pequeño responderá siempre y cuando sea suficientemente fuerte para calentar un alambre metálico al rojo vivo.

Los extremos opuestos de la batería galvánica estaban unidos por un alambre metálico, el cual, para resumir llamaremos el *conductor de unión*, o el *alambre de unión.* Al efecto que tiene lugar en este conductor y en el espacio circundante, daremos el nombre de *conflicto de la electricidad*.

Dejé que la parte recta de este hilo se colocara horizontalmente por encima de la aguja magnética, debidamente suspendida, y paralela. Si es necesario, el alambre de unión se dobla para asumir una posición apropiada para el experimento. Estando las cosas en este estado, la aguja se moverá, y el extremo que está cercano al lado negativo de la batería irá hacia el oeste.

Si la distancia del alambre de unión no excede tres cuartos de una pulgada desde la aguja, la declinación de la aguja forma un ángulo de cerca de 45°. Si la distancia se incrementa, el ángulo disminuye proporcionalmente. La declinación asimismo varía con la potencia de la batería.

El alambre de unión puede cambiar de lugar, ya sea hacia el este o hacia el oeste, siempre y cuando continúe paralelo a la aguja, sin ningún otro cambio del efecto que no sea en relación a su cantidad. Por lo tanto, el efecto no puede atribuirse a la atracción; porque el mismo polo de la aguja magnética, que se acerca al alambre de unión, mientras está situado en su lado este, debería retroceder cuando está en el lado oeste, si estas declinaciones dependen de atracciones y repulsiones. El conductor de unión puede consistir en varios alambres, o bandas metálicas, conectadas entre sí. La naturaleza del metal no altera el efecto, sino simplemente la cantidad. Alambres de platino, oro, plata, latón, hierro, cintas de plomo y estaño, una masa de mercurio fueron empleados con igual éxito. El conductor no pierde su efecto, aún interrumpido por agua, a menos que la interrupción llegue a ser de varios centímetros de longitud.

El efecto del alambre de unión pasa a la aguja a través de vidrio, metales, madera, agua, resina, cerámica, piedras; ya que este no se elimina por interposición de placas de vidrio, metal o madera. Incluso el vidrio, el metal y la madera, interpuestos al mismo tiempo, no destruyen, y ciertamente apenas disminuyen el efecto. El disco del electróforo, placas de porcelana, un recipiente de piedra, incluso lleno de agua, se interpusieron con el mismo resultado. Encontramos los mismos efectos cuando la aguja estaba dentro de una caja de latón llena de agua. Es innecesario observar que la transmisión de efectos a través de todos estos materiales nunca antes se ha observado en electricidad y galvanismo. Los efectos, por lo tanto, que tienen lugar en el conflicto de electricidad son muy diferentes de los efectos de cualquiera de las electricidades.

Si el alambre de unión se coloca en un plano horizontal bajo la aguja magnética, todos los efectos son los mismos que cuando está por encima de la aguja, sólo que están en una dirección opuesta; para el polo de la aguja magnética próximo al extremo negativo de la batería se tiene que se desvía al este.

Para que estos hechos sean más fáciles de retener, podemos usar esta fórmula- el polo superior por el cual entra la electricidad negativa se gira hacia el *oeste*; el inferior, al *este*.

Si el alambre de unión se gira en un plano horizontal para formar un ángulo gradualmente creciente con el meridiano magnético, la declinación de la aguja *aumenta*, si el movimiento del alambre es hacia el lugar de la aguja perturbada; pero d*isminuye* si el alambre se mueve más lejos de ese lugar.

Cuando el alambre de unión está situado en el mismo plano horizontal en el que se mueve la aguja por medio del contrapeso, y paralelo a él, no se produce ninguna declinación ni al este ni al oeste; pero tiene lugar una inclinación, de modo que el polo, al lado del cual la electricidad negativa entra en el alambre, se *deprime* cuando el alambre está situado en el lado *oeste*, y se *eleva* cuando está situado en el lado *este*.

Si el alambre de unión se coloca perpendicularmente al plano del meridiano magnético, ya sea por encima o por debajo de él, la aguja permanece en reposo, a menos que esté muy cerca del polo; en ese caso el polo se *eleva* cuando la entrada es del lado *oeste* del alambre, y *deprimido*, cuando es del lado *este*.

Cuando el alambre de unión se coloca perpendicularmente opuesto al polo de la aguja magnética, y el extremo superior del alambre recibe la electricidad negativa, el polo se mueve hacia el este; pero cuando el alambre es opuesto a un punto entre el polo y el centro de la aguja, el polo está más hacia el oeste. Cuando el extremo superior del alambre recibe electricidad positiva, los fenómenos se invierten.

Si el alambre de unión se dobla de modo que forme dos secciones paralelas entre sí, repele o atrae los polos magnéticos de acuerdo a las diferentes condiciones del caso. Supongamos que el alambre se coloca opuesto a cualquiera de los polos de la aguja, de modo que el plano de las secciones paralelas es perpendicular al meridiano magnético, y dejamos la sección oriental unida con el extremo negativo, la sección occidental con el extremo positivo de la batería: en ese caso el polo más cercano será repelido hacia el este o hacia el oeste, según la posición del plano de las secciones. Estando unida con el positivo la sección que está más al este, y la que está más al oeste con el lado negativo de la batería, el polo más cercano será atraído. Cuando el plano de las secciones se coloca perpendicular al lugar entre el polo y el centro de la aguja, se repiten los mismos efectos, pero invertidos.

Una aguja de latón, suspendida como una aguja magnética, no se mueve por el efecto del alambre de unión. Igualmente las agujas de vidrio y de goma de laca permanecen inactivas.

Podemos ahora hacer algunas observaciones para explicar estos fenómenos.

El conflicto eléctrico actúa sólo sobre las partículas magnéticas de la materia. Todos los cuerpos no magnéticos parecen penetrables por el conflicto eléctrico, mientras que los cuerpos magnéticos, o más bien sus partículas magnéticas, resisten el paso de este conflicto. Por lo tanto pueden ser movidos por el impulso de las potencias contendientes.

Es suficientemente evidente por los hechos anteriores que el conflicto eléctrico no está confinado al conductor, sino que se dispersa ampliamente en el espacio circundante.

De los hechos precedentes podemos igualmente colegir que este conflicto realiza círculos; porque sin esta condición, parece imposible que la parte del alambre de unión, cuando es colocada debajo del polo magnético, deba conducirlo hacia el este, y cuando es colocada encima deba conducirlo hacia el oeste; porque es la naturaleza de un círculo que los movimientos en partes opuestas deben tener una dirección opuesta. Además, un movimiento en círculos, junto con un movimiento progresivo, de acuerdo a la longitud del conductor, debe formar una línea concoidal o espiral; pero esto, a menos que me equivoque, no contribuye a la explicación de los fenómenos hasta ahora observados.

Todos los efectos sobre el polo norte magnético arriba mencionados se entienden fácilmente suponiendo que la electricidad negativa se mueve en una línea espiral curvada hacia la derecha, e impulsa el polo norte, pero no actúa sobre el polo sur. Los efectos sobre el polo sur se explican de manera similar, si atribuimos a la electricidad positiva un movimiento contrario y el poder de actuar en el polo sur, pero no sobre el norte. El acuerdo de esta ley con la naturaleza será mejor visto por una repetición de los experimentos que por una larga explicación. El modo de juzgar de los experimentos se facilitará mucho si el curso de las electricidades en el alambre de unión se señala por marcas o figuras.

Simplemente añadiré a lo anterior que he demostrado en un libro publicado hace cinco años que el calor y la luz consisten en un conflicto de electricidades. De las observaciones ahora expuestas, igualmente podemos concluir que un movimiento circular ocurre en estos efectos. Creo que esto contribuirá mucho a ilustrar los fenómenos a los que se ha dado la denominación de polarización de la luz.

Copenhague, julio 21, 1820

4. Segundo escrito (noviembre de 1820)**

Nuevos experimentos electromagnéticos

Desde la publicación de mis primeros experimentos sobre la acción magnética de la batería galvánica, he multiplicado mis investigaciones sobre ese tema tanto como una multitud de otras ocupaciones importantes lo pusieron en mi poder.

Los efectos magnéticos no parecen depender de la intensidad de la electricidad, sino únicamente de su cantidad. La descarga de una batería eléctrica fuerte transmitida a través de un cable metálico no produce alteración en la posición de la aguja magnética. Una serie de chispas eléctricas interrumpidas actúa sobre la aguja por las atracciones y repulsiones eléctricas ordinarias, pero hasta donde se puede percibir, las chispas no producen ningún efecto electromagnético. Una pila galvánica compuesta de 100 discos de metal de dos pulgadas cuadradas, y de papel humedecido con agua salada para que sirva de conductor de fluido, carece igualmente de efecto sensible sobre la aguja. Por otro lado se obtiene el efecto con un solo arco galvánico de zinc y cobre que tiene por conductor un líquido de gran poder conductor; por ejemplo, de una parte de ácido sulfúrico, otra de ácido nítrico y 60 partes de agua. Incluso podemos duplicar la cantidad de agua sin disminuir mucho el efecto. Si la superficie de los dos metales es pequeña, el efecto también es pequeño. Pero aumenta en la proporción en que aumentamos la superficie. Una placa de zinc, de seis pulgadas cuadradas, sumergida en un recipiente de cobre que contiene el líquido conductor del que he hablado, produce un efecto considerable. Pero una disposición de este tipo en la que la placa de zinc tiene una superficie de 100 pulgadas cuadradas actúa sobre la aguja con tal fuerza que el efecto es muy sensible a una distancia de tres pies, incluso cuando la aguja no es muy móvil. No he observado mayores efectos de un aparato galvánico compuesto por 40 canales similares; de hecho, el efecto parecía menos grande. Si esta observación, que no he investigado expresamente, es justa, tendré la opinión de que la pequeña disminución de la potencia conductora producida al aumentar el número de elementos del aparato debilita su efecto electroquímico.

Para comparar el efecto de un solo arco galvánico con el de un aparato compuesto por varios arcos o elementos, haremos una observación. La fig. 9 [figura 1], representa un arco galvánico compuesto por una pieza de zinc *z*, de cobre *c*, de un alambre metálico *a b* y

de un líquido conductor *l*. El zinc siempre comunica una parte de su electricidad positiva al agua como lo hace el cobre con su electricidad negativa. Esto ocasionaría una acumulación de electricidad negativa en la parte superior del zinc y de electricidad positiva en la parte superior del cobre, a menos que la comunicación *a b* restableciera el equilibrio al permitir un paso libre para la electricidad negativa de *z* a *c* y para la electricidad positiva de *c* a *z*. Vemos entonces que el alambre *a b* recibe la electricidad negativa del zinc, y la electricidad positiva del cobre, mientras que un alambre que constituye la comunicación de los dos polos de una pila, o de otro aparato galvánico compuesto, recibe la electricidad positiva del polo de zinc y la electricidad negativa del polo de cobre.

Prestando atención a esta distinción, podemos, con un solo arco galvánico, repetir todos los experimentos que hice al principio con un aparato galvánico compuesto. Emplear un solo arco galvánico da esta gran ventaja, que nos permite repetir los experimentos con poca preparación y gasto. Pero presenta otra ventaja aún más considerable; es decir, que podamos establecer un arco galvánico suficientemente potente para los experimentos electromagnéticos y, sin embargo, suficientemente ligero para ser suspendido a un pequeño alambre metálico, de tal manera que el pequeño aparato pueda hacerse girar alrededor del eje prolongado del alambre. De esta manera podemos examinar la acción que ejerce un imán sobre el arco galvánico. Como un cuerpo no puede poner otro en movimiento sin ser movido a su vez, cuando posee la movilidad requerida, es fácil prever que el arco galvánico debe ser movido por el imán.

Utilicé diferentes disposiciones del aparato galvánico simple para examinar el movimiento que le imprimía el imán. Uno de estos arreglos está representado en la fig. 10 [figura 1], que representa una sección perpendicular del mismo en la dirección del ancho. *c c c c* es una artesa de cobre, de tres pulgadas de alto, cuatro pulgadas de largo y media pulgada de ancho. Sin duda, estas dimensiones pueden variar hasta el infinito. Solo es

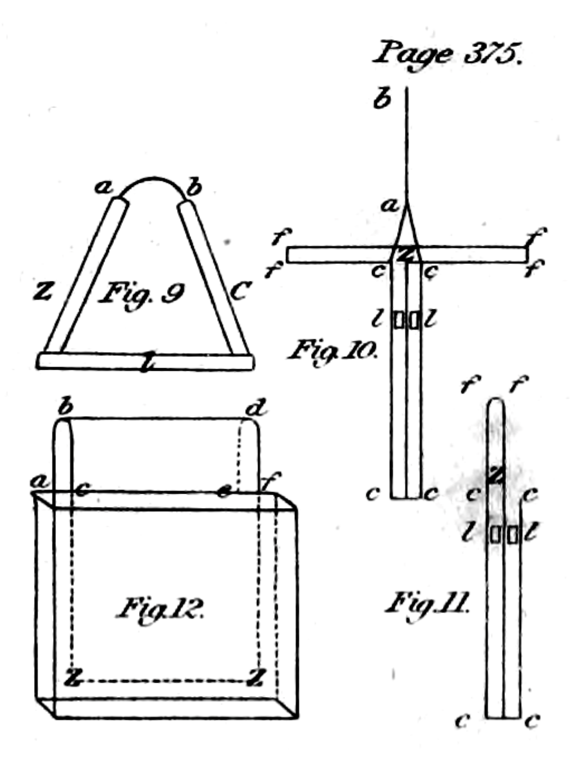


Figura 1. Disposiciones experimentales con arcos galvánicos. Fuente: Ørsted, H. (1820). New electromagnetic experiments, *Annals of Philosophy, 16*, placa CIX, figuras 9-12, p. 352.

necesario observar que la anchura no debe ser grande, por lo que la artesa debe estar hecha de placas lo más delgadas posible. *z z* es una placa de zinc. *l l* son dos piezas de corcho que mantienen las placas en su posición, *c f f f f z* es un alambre de latón, de un cuarto de tono al menos de diámetro, *a b* es un alambre de latón lo más fino posible, para poder soportar el peso del aparato, *c a c* es un hilo de lino que une el alambre al aparato. La cubeta contiene el líquido conductor. El hilo conductor de este aparato atraerá el polo norte de la aguja cuando se coloque en el lado izquierdo del plano *c f f f f z*, considerado en la dirección *f z*. En el mismo lado será repelido el polo sur. En el otro lado de este plano, el polo norte será repelido y el polo sur atraído. Para que se produzca este efecto, no debemos colocar la aguja por encima de *f f*, ni por debajo de *f z* o *f c*. Si en lugar de acercar una pequeña aguja móvil al hilo conductor acercamos una de las extremidades de uno de los polos de un imán energético, la atracción o repulsión indicada por la aguja, pondrá en movimiento el aparato galvánico, y lo hará girar alrededor del eje prolongado de *a b*.

Si en lugar del hilo conductor tomamos una gran cinta de cobre del mismo ancho que la placa de zinc, el efecto difiere del que acabamos de mencionar sólo en que es mucho más débil. Por otro lado, aumentamos un poco el efecto haciendo que el conductor sea muy corto. La figura 11 [figura 1], representa la sección perpendicular de esta disposición en la dirección de la anchura de la artesa. La figura 12 [figura 1], muestra la misma disposición en perspectiva. Es obvio que *a c b d e f* representa la placa conductora y *c z z f* la placa de zinc. En esta disposición, el polo norte de la aguja será atraído hacia el plano de *a b c*, y el polo sur será repelido desde el mismo plano. *e d f* tendrá efectos contrarios. Aquí tenemos un aparato cuyas extremidades actúan como los polos de la aguja. Pero hay que reconocer que solo las caras de las dos extremidades, y no las partes intermedias, tienen esta analogía.

Asimismo, podemos hacer un aparato galvánico móvil de dos placas, una de cobre y otra de zinc, retorcidas en espiral y suspendidas en el conductor de fluido. Este aparato es más móvil; pero son necesarias más precauciones para no dejarse engañar cuando hacemos experimentos con él.

Todavía no he encontrado un método para hacer un aparato galvánico capaz de dirigirse hacia los polos de la tierra. Para este objeto sería necesario poseer un aparato mucho más móvil.

5. Conclusiones

El camino recorrido a través del experimento de la aguja permite visualizar varios hechos. En primer lugar, genera un cuestionamiento sobre el abordaje que se ha realizado sobre los aportes experimentales y teóricos de Ørsted, quien según lo mencionan Jackson, Jelved y Knudsen (1998), también realizó investigaciones sobre la compresibilidad de los gases y la producción de compuestos químicos como cloruro de silicio, no obstante, la publicación de estos trabajos en diferentes idiomas como danés, alemán, latín, francés o inglés, supuso una gran dificultad para los académicos que pretendían conocer su vida y obra.

En segundo lugar, los documentos dispuestos anteriormente conducen a un cuestionamiento sobre lo sucedido en la vida académica de Ørsted posterior al experimento de la aguja. Al respecto, es posible establecer que dedicó un tiempo a la exposición de su explicación sobre el conflicto eléctrico en Europa sin tener mayor éxito, pues por una parte, en Londres, Faraday no entendía la teoría en tanto parecía vaga y difusa, mientras que en París, los seguidores de Ampere llegaron a poner en tela de juicio la autenticidad de su teoría, señalando que era derivada de la teoría electrodinámica.

En tercer lugar, el experimento permite contemplar el impacto que tuvo la experiencia a nivel teórico y experimental en la comunidad científica, en tanto, supuso diferentes experiencias con el propósito de caracterizar y diferenciar los fenómenos electrostáticos de los electrodinámicos (como aquellas desarrolladas por Ampere), y estos en su conjunto, de los fenómenos magnéticos, lo cual favorecería tener una claridad de los efectos (mecánicos, térmicos, eléctricos y magnéticos en los cuerpos) y un análisis de sus causas, situación que en años posteriores, se manifestaría en una convergencia en el lenguaje empleado. Así mismo, el experimento de Ørsted da la posibilidad de contemplar la inquietud que se tenía en el siglo XIX sobre la naturaleza de la electricidad y el magnetismo, y se dispone como una invitación a analizar las diferentes maneras de explicar los fenómenos electromagnéticos, las cuales en principio estaban sometidas a una fuerte influencia de la herencia newtoniana, no obstante, la transversalidad y la no instantaneidad de la acción se constituyeron en una evidencia posterior que llevó a repensar la acción experimentada por un cuerpo, entendiendo que era consecuencia de variaciones del estado del espacio vecino.

De esta manera, tales formulaciones se distancian de la postura de Newton al tratar de dar cuenta de las diferentes situaciones en términos de atracciones y repulsiones dependientes de la masa de los cuerpos, permitiendo así, desarrollar un conjunto de explicaciones sobre los fenómenos electromagnéticos al involucrar otros elementos y concepciones que de paso, se constituyen en las bases de la idea de campo que posteriormente, sería refinada con el trabajo de James Clerk Maxwell.

En cuarto lugar, el experimento de la aguja de Ørsted favorece una nueva mirada sobre los fenómenos científicos pues el hombre instituye técnicas que no parten de un examen empírico de la naturaleza, y su aplicación exige la organización de la experiencia y las observaciones (particularmente en el campo del electromagnetismo), para poder enriquecer el marco experimental y conceptual al develar fenómenos que no se ofrecen al examen inmediato (Bachelard, 1978), es decir, los fenómenos no se presentan de manera espontánea, sino que se convierten en entidades que se fabrican en la interacción entre los instrumentos y las respuestas de la naturaleza (Romero, 2013).

Por último, el abordaje de la historia de las ciencias favorece la comprensión de discursos, la adquisición de herramientas para determinar por qué una proposición se estima justificada (Matthews, 1994), y la identificación de posturas que toman los científicos dentro de un marco intelectual, de tal forma que se tiene una mayor comprensión de los contenidos de las ciencias, en términos de su lógica, así como sus limitaciones y las cosmovisiones que los subyacen, y de los elementos que en algún momento los determinaron. Así pues, la revisión de la historicidad de las ciencias favorece una visión de las mismas como una actividad que se desenvuelve en contextos socio-culturales específicos, en donde se construyen explicaciones de fenómenos del mundo que describen un camino hasta constituirse en hechos científicos socialmente aceptados. De tal forma, la ciencia no se trataría de productos con carácter objetivo e independiente al sujeto, sino construcciones provisionales y atemporales (Rodríguez & Romero, 1999).

Referencias

Andrade, R. (2001). Romagnosi and Volta's Pile: Early difficulties in the interpretation of voltaic electricity. En F. Bevilacqua & L. Fregonese (Ed.), *Nuova Voltiana: Studies on Volta and his times 3* (pp. 81–102). Pavía; Milán: Ulrico Hoepli Editore.

Bachelard, G. (1978). *El Racionalismo aplicado.* Buenos Aires: Paidos.

Berkson, W. (1981). *Las teorías de los campos de fuerza: Desde Faraday hasta Einstein.* Madrid: Alianza Editorial.

Brain, R., Cohen, R., & Knudsen, O. (2007). *Hans Christian Ørsted and the Romantic Legacy in Science: Ideas, Disciplines, Practices.* New York: Springer.

Faraday, M. (1821). Historical Sketch of Electromagnetism. *Annals of Philosophy*, *2*(3), 195–199.

Jackson, A., Jelved, A. & Knudsen, O. (Eds). (1998). *Selected scientific Works of Hans Christian Orsted.* New Jersey: Princeton University Press.

Kuhn, T. (2004). *La estructura de las revoluciones científicas.* México: Fondo de Cultura Económica. (Obra original de 1962)

Matthews, M. (1994). Historia, filosofía y enseñanza de las ciencias: la aproximación actual. *Enseñanza de las ciencias, 12* (2), 255–277.

Millikan, R. (1944). *Electrones (+ y -), protones, fotones, neutrones y rayos cósmicos.* Buenos Aires-México: Editorial Espasa-Calpe Argentina S.A.

Ørsted, H. (1820). Experiments on the effect of a current of electricity on the magnetic needle. *Annals of Philosophy*, *16*, 273–276.

Ørsted, H. (1820). New electromagnetic experiments. *Annals of Philosophy*, *16*, 375-377.

Rodríguez, L., & Romero, A. (1999). La construcción de la historicidad de las ciencias y la transformación de las prácticas pedagógicas. *Física y Cultura: Cuadernos sobre historia y enseñanza de las ciencias*, *6*, 3–16.

Romero, A. (2013). La experimentación como potenciadora de reflexiones sobre la naturaleza de las ciencias. En M. Ayala, J. Malagón, & S. Sandoval (Eds.), *Construcción de fenomenologías y procesos de formalización: un sentido para la enseñanza de las ciencias* (pp. 37–54). Bogotá, Colombia: Universidad Pedagógica Nacional.

Declaraciones

Conflictos de interés: El autor declara que no existen conflictos de interés.